MASS. ED21.2: B46 ✓

BEYOND Paper & Pencil

* UMASS/AMHERST *
312066 0271 3146 1

Background Summary

ERNMENT DOCUMENTS
COLLECTION
1 - 1993
University of Massachusetts

What Is Performance Testing?

Performance testing is usually associated with some physical criteria—the speed and efficiency of a racing car, the skill and agility of a downhill racer. Consequently,it may seem somewhat counter-intuitive to speak of it in relation to problem solving, especially when speed is not considered to be a factor. The term is used in this program to highlight the fact that the problems are set in a practical context and that they require physical, as well as intellectual, activity. It is also used to indicate the focus of evaluation. In contrast to the majority of written tests, the criteria for success on performance tests are concerned less with the solution to the problem than with the strategies in achieving that solution. The premise of these tests is that students' actions in solving problems reveal much about their understanding of the subject and their ability to reason well. Primarily, it is intellectual, not physical, performance that we are measuring in these tests.

Why Performance Testing?

Performance testing has important implications for both students and their teachers.

It is important for students because it demonstrates what we value in education. Research confirms what teachers have known for a long time—that students learn best what they are tested on. It follows, then, that if we rely solely upon simple tests of discrete skills and information, we should not be surprised to find our students performing poorly when asked to go beyond the facts, to think for themselves. However, if we value the generation of ideas, the application of principles, the active involvement of our students, we should be testing for these qualities. A short answer test that demands only the recitation of known facts sends a strong message to the contrary. Performance testing, on the other hand, gives a message to students about the kind of active problem solving that we expect as an outcome of schooling.

This type of testing is important for teachers in terms of their own teaching and in terms of the validity of their evaluations.

Increasingly, we are beginning to recognize that students carry to school a full and complex set of notions about the world. Ideally, school learning interacts with prior knowledge to forge new and more complete understandings. How-

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education January 1990

In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

924/172

ever, this is often not the case in practice. For many children, school learning becomes firmly bound to its classroom context. Seldom do they see the connection between what they learn in science and mathematics and the practical problems that make up their everyday world. In short, they do not transfer their learning.

Performance testing allows teachers to judge how well students have made these connections and to what extent they can transfer what they learn in school to the world about them.

Performance testing can also be used to increase the validity of teachers' evaluations of student achievement Grades that depend upon only one type of information are apt to give a distorted picture of what students know. A valid assessment has to incorporate information gathered from different sources, showing different kinds of knowing. That students know the rules is not sufficient evidence to make valid judgements about their understanding. Students should be also judged on their ability to solve problems that call upon them to integrate and apply their knowledge and skills.

How Performance Testing Fits in the Massachusetts Educational Assessment Program

Chapter 188 of the Act of 1985 established the Massachusetts Educational Assessment Program in order to compare the effectiveness of public schools and to give guidance for the improvement of curriculum and instruction. Because of this dual intent, the program has developed three major components which satisfy different audiences.

The Multiple-Choice Assessment: For the public at large, the most prominent part of the assessment program is the multiple-choice component, which is administered in alternate years to all fourth, eighth and twelfth grade students throughout the Commonwealth. Covering the areas of reading, mathematics, science and social studies, over 3500 items are used to measure a wide range of content and types of thinking appropriate to these subjects. Since the purpose of the test is to report on school rather than on individual achievement, each student takes only a portion of the complete test. Results are reported to each school and district throughout the Commonwealth and are a matter of public record. Using this information, schools and districts can review the overall effectiveness of their program and its effectiveness for specific groups of students.

The Open-Ended Assessment: A second component of the assessment consists of the administration of open-ended questions in reading, mathematics, science and social studies. These questions were administered to a sample of students as part of the multiple-choice assessment; however, they are seen as a separate entity. They do not form part of the school reports and are analyzed separately by groups of subject area experts. As a result, the questions in the open-ended assessment focus on students' beliefs and understanding, as well as their ability to apply knowledge in less traditional contexts. They demand the kinds of thinking that cannot be captured by an item that asks only for the identification of a correct option. The results of the 1988 open-ended assessment have been published in a series of reports entitled **On Their Own.** These reports contain, not only analyses of the results and their instructional implications, but the questions themselves. Classroom teachers are encouraged to administer them to their students and to make normative comparisons if they so desire.

The Performance Assessment: In 1989, the Department expanded further the dimensions of the testing program by a third component to the program. Loosely called "performance testing," it involves the application of mathematical and scientific concepts to the solutions of problems. These tests, administered in alternate years to the written assessment, ask fourth and eighth grade students to apply what they know about mathematics and science to solve problems. Since they are designed to measure how well students are able to integrate their knowledge, each test calls for a number of different concepts, as well as skills in measuring, calculating, observation, etc.

The results of these tests are reported in a series of brochures entitled **Beyond Paper and Pencil.** In each brochure, the tests themselves are published, along with analyses, instructional implications, and suggestions for extension to classroom instruction and evaluation.

The three testing formats—multiple-choice, open-ended and performance—are designed to provide a rounded picture of what students know, how they apply that knowledge, and how they assimilate knowledge to carry out multi-

dimensional tasks. While the results of the multiple-choice are used to calculate the schools scores which are used for comparison purposes, the results of the open-ended and performance tests are reported in order to improve curriculum and instruction. They are intended to be used by principals, teachers, and curriculum coordinators who are primarily interested in the achievement of individual students. Not only are they intended to broaden teachers' understanding of students' conception and difficulties, but they are presented as models for more valid evaluation of individual students.

The Rationale and Development of the Performance Tests

The prime concern in the choice of performance tasks was their relevance to the mathematics and science curriculum. Within these areas, there is general consensus that students should be able to "act like scientists" or "act like mathematicians." On the other hand, in real life, scientific and mathematical skills have little value in themselves; they are tools to be used in the context of a task. Measuring, weighing, calculating and hypothesizing are only of interest because of their value in solving problems or coming to a solution. For this reason, we looked for tasks that measured both procedures and understanding. Each task was presented within the context of some basic area of science or mathematics and was designed to measure students' understanding of the concepts involved, as well as their ability to use the appropriate skills.

A second concern was to develop tasks that reflected the social aspect of learning. Whereas most written tests are essentially solitary endeavors, real problems are usually solved cooperatively with a give-and-take of ideas. Not only are such tasks more reflective of the kinds of problems that are posed in everyday life but they are more amenable to observation. When students communicate with each other, defending or explaining their thoughts, inferences about their ideas are more apt to be valid than when one only infers from a solution or single response.

A third concern was to establish reliability among the various administrators. One way of establishing reliability is to turn performance testing into a behavioral form of multiple-choice—that is, to limit the administrators' recordings to a narrow set of observable behavior (e.g., "measured temperature accurately"). Another way is to include as much behavior and dialogue as possible, relying upon content analysis to establish uniform trends among the testing situations. We chose to combine both approaches, asking the administrator 1) to record students' competencies and methods in dealing with the task which has been set, and 2) to record their own observations of students' conversations and strategies as completely as possible. In addition, administrators were asked to discuss the task with the students upon completion if clarification was needed.

As a result of these considerations the tests were developed and piloted during the autumn and winter of 1988. At the same time, video tapes were made as a record of the tests and for training purposes. Detailed scoring scripts were also developed at this time.

During the month of November 1988, a brochure was mailed to all schools and members of science and mathematics organizations describing plans for performance testing and asking for volunteer administrators. Approximately 135 teachers and curriculum coordinators responded. From that group, 65 were chosen to be trained. Training was conducted during a two-day residential conference during which administrators worked with the tests, watched video tapes of students completing the tasks and practiced scoring from the tapes. In addition, some changes were made in the scoring guides and scripts according to suggestions from the administrators. Administrators then returned to their schools and practiced giving the tests to their own students. Their scripts were corrected by Department staff in order to ensure standardized responses.

Administration

In January 1989, 150 fourth and 75 eighth grade schools were chosen to participate in the testing program. The schools were selected according to a stratified random sampling procedure which accounted for geographical diversity as well as relative performance on the science portion of the multiple-choice assessment.

The testing took place in two phases. The first phase, the administration of a short (30-minute) multiple-choice mathematics and science test, was conducted during the week of February 6. This written test measured many of the concepts found in the performance tests and served as the basis for selection of a sample of students, representing all levels of ability as measured by

Circuits (Fourth and Eighth Grades)

Description of the Task: (1) Given a battery, wires, and light bulb, students were asked to light a bulb. (2) They were asked to light a bulb using a circuit board and to describe the flow of electricity. (3) They were presented with four identical boxes and asked to use a set of materials to identify which contained a battery, a wire, a toothpick and a resistor.

Concepts/Skills Evaluated:

- electrical circuitry
- use of apparatus
- formulation of hypotheses
- accurate observation of phenomena
- analysis and interpretation of data
- drawing conclusions
- methodical and efficient working

Insulation (Eighth Grade)

Description of the Task: Students were asked to evaluate the insulating capacity of three hot drink cups and to come to a decision based on their findings. They were provided with the materials necessary to carry out and record such an investigation.

Concepts/Skills Evaluated:

- analysis of the experimental situation
- identification and control of variables
- selection of appropriate procedures and apparatus
- performance of specific techniques to ensure control
- accurate observation of objects and phenomena
- skills: reading instruments, recording
- analysis and interpretation of data

Popcorn Estimation (Fourth and Eighth Grades)

Description of the Task: Students were told to use whatever they wanted from a set of materials to estimate the number of kernels in a container. The administrator observed and recorded their strategy. When they had reached a solution, they were asked to describe an alternate method and, if appropriate, to carry it out.

Concepts/Skills Evaluated:

- identification of problem
- appropriate procedures
- correct use of apparatus (balance)
- proportional reasoning
- computation
- understanding of large numbers
- accuracy/precision judgments

Cubes (Fourth and Eighth Grades)

Description of the Task: Students were asked to: (1) construct and record the set of rectangular solids that could be made with 16 cubes; (2) use a small number of cubes to find the total number of cubes that would fill a container; and (3) calculate the number of cubes needed to complete each of three incomplete rectangular solids of increasing complexity.

Concepts/Skills Evaluated:

- concept of volume
- factorial relationships
- spatial visualization
- computational skills
- strategies

Animal/Leafkey (Fourth and Eighth Grades)

Description of the Task: (1) Students were shown cards and asked to identify which illustrated an animal and to classify the remainder. (2) They were asked to sort a set of clown faces and to describe their categorizations. (3) They were asked to sort a set of leaves and to describe their classifications.

Concepts/Skills Evaluated:

- categories of living things: concept of animal
- categorization according to attributes
- general schemata vs. arbitrary classification
- verbal ability

Playground (Fourth Grade)/Math Town (Eighth Grade)

Description of the Task: Given maps of the various areas, students were asked to place objects according to a set of constraints. They were then given a series of tasks that involved measurement of area and perimeter.

Concepts/Skills Evaluated:

- organizational ability
- dealing with logical and practical constraints
- mathematical concepts: area, perimeter, locus, scale
- skills: measuring, computation, use of instruments

the test, as well as forming a useful basis for comparison with the performance testing.

In March each school received a list of selected students, as well as information regarding the test administration. Suggestions regarding how and when to inform the selected students and other issues were dealt with in a memorandum accompanying the list of chosen students. Schools were also given a sample letter which explained the testing and could be sent to parents if desired.

Performance testing took place between March 20 and April 14, 1989. Each elementary school was visited once by a test administrator; each eighth grade school was visited on two occasions. On the day of the testing, the administrator proceeded through the list of selected students who, in turn, were asked to choose a partner to join her in completing the task. A single problem was administered to each pair of student. Since the tasks varied in length and the students themselves varied in the amount of time taken for solution, testers were told to expect between 6 and 10 pairs of students within a single day. This proved to be an accurate estimate.

Sixty-three administrators visited between two and five different schools over a 5-day period. As a result of the testing, approximately 2000 selected students and their partners completed one of the performance tasks. Their behavior was carefully observed and recorded, then analyzed by the Department of Education.

Test Results

The seven tasks ranged in structure from extremely structured to completely open-ended. Four dealt primarily with mathematics concepts and processes; three with science. **Insulation** was given at the eighth grade level only; while **Math Town** and **Playground** were parallel tasks, given at the eighth and fourth grades, respectively. A short description of each task with its focus is given on the centerfold. More complete descriptions—including results, examples and extensions for classroom instruction/evaluation—are contained in separate brochures. The summary below discusses only the major findings.

Popcorn Estimation, the most open-ended of all the tasks, was clearly unfamiliar to many students and proved extremely difficult for both fourth and eighth grade students. The task evoked a number of possible strategies—from the rather straight-forward but relatively inefficient method of counting the number of kernels in a small container and then using that container to measure all the kernels to the more sophisticated method of determining equivalencies between different measures (e.g., grams or milliliters).

In addition to a lack of familiarity with using the apparatus (31 percent of fourth graders and 39 percent of eighth graders were observed to add weights to the heavier side of the balance or to remove them from the lighter side), observations of students' strategies showed a lack of understanding of the relationship between different units of measurement. They appeared to lack an understanding of equivalency—not only physical equivalency (as in the case of the balance) but in terms of the relationship between different units of measurement. The requirement of reasonableness in estimation also presented some problems. As the balance so attracted some students that they suspended all judgements concerning its actual use, numerical computation appeared to obscure the judgement of others. These students shown an uncritical reliance on their computation, with little checking against reality.

In contrast, **Animal/Leafkey** was a very structured task, examining students' language and concepts in addition to their ability to classify. It highlighted the crucial role that language plays in thinking and understanding in science. Although most students readily identified land animals as within the category of animal, more experienced difficulty with less typical examples such as fish, bumblebees, and snails. Only 26 percent of fourth graders and 55 percent of eighth graders correctly identified a boy as an animal. When asked to classify a set of figures that was designed to display the presence or absence of certain attributes, students at both grade levels had few problems. However, when presented with a less artificial grouping (a set of leaves), they experienced difficulty in articulating the characteristics that could be used to subdivide the set.

Like **Popcorn Estimation**, **Insulation** was deceptively simple. Administered only at the eighth grade level, this task represented the only actual experiment that students were asked to perform. Few students had clear idea of what the task involved, trying only to find the cup in which the temperature of the water was the highest. It is doubtful that many recognized it as a scientific experiment or saw the need for identifying and controlling variables. Most

were concerned only with producing an answer. Even here students did not relate their answer to the practical requirements of the problem. For example, only 27 percent took readings over a period of 4 minutes or longer. Although half recorded the results, little attempt was made to use the data. Few made any attempt to examine it for patterns or anomalies such as different rates of heat loss. They interpreted the task as one that called for producing an answer rather than examining a relationship.

Although much of **Cubes** was devoted to the use of spatial visualization in determining volume, its initial problem involved a concrete representation of numerical relationships. Few students recognized this. Most concentrated on making discrete figures with the 16 cubes provided without recognizing the numerical pattern that emerged.

Unfamiliarity with the concept of volume undoubtedly accounted for the large differences between fourth and eighth graders in the other problems in this task. Although half of the fourth graders could give volume estimates when presented with a box and cubes, the additional requirement that they use an incomplete solid to visualize a completed one was extremely difficult for them. Only a quarter could complete the simplest figure presented. Eighth graders also found it difficult to compute the number of missing cubes in an incomplete rectangular solid, particularly as the solids became more complex and more dependent upon visualization strategies. Gender differences were noted between girls and boys in strategies, although not in success in execution.

Although **Circuits** dealt with the essential scientific processes of confirming or disconfirming hypotheses through observation, it was structured in such a way that most of the organizational demands were eliminated. A chart and a series of leading questions helped students to focus on the main problems of prediction and observation. Although only 42 percent of fourth graders and 60 percent of the eighth graders showed an understanding of the flow of electricity in a circuit, this did not interfere with their practical understanding of how to use batteries and wires to light a bulb. Almost all students at both grades were able to complete this task; while, approximately 28 percent of fourth graders and 45 percent of eighth graders correctly identified the contents of all the mystery boxes. When asked to generate and test their hypotheses in order to do this, those who had some knowledge of conductivity were at an advantage. Students who recognized that they *did not know* also managed to come to a correct solution by a process of elimination. However, those who had misconceptions about the effect of different materials found it difficult to confirm or disconfirm their hypotheses purely by observation.

Although students were fairly successful in dealing with the constraints required by **Playground** and **Math Town**, they dealt less well with the measuring tasks involved. With the exception of the ruler, students showed a reluctance to use the measuring instruments provided, even in cases where they were clearly more appropriate. This was true of both grade levels, The added restrictions at grade eight resulted in a more difficult task for students at this level, with a relatively small percentage of students meeting all the restrictions. In both tasks, students had to organize their responses to various requirements while remaining flexible enough to adjust to different restrictions. The area determination problems proved very difficult at both grades; however, fourth graders seemed more inclined to use unconventional though appropriate means (i.e., the grid) to find the answer. Eighth graders seemed to believe that calculation was the only means of finding the area even though they had been provided with a grid. For eighth graders, because one of the areas was an irregular shape, linear measurement was clearly inappropriate.

The least structured tasks proved to be the most revealing of students' understanding, as well as of their strategies. However, the main impression gained from these open-ended tasks was of students' unfamiliarity with the process of investigation. Despite the administrators' request that they "form a plan," most students at both level started **Popcorn Estimation** without any clear analysis of what was involved, only to find that they later had to change strategies. Again, for many at both levels, the sheer quantity of the kernels and the availability of different measures led to confusion and many lost sight of the initial problem. **Insulation** led to some of the same behavior. Students did not appear to see the parallel between this practical problem and their experiences in science labs. Many took the question at face value. In neither case did they ask clarifying questions or attempt to analyze the requirements of the task.

In addition, students did not make appropriate use of the equipment available. In **Popcorn Estimation**, the presence of the balance was extremely attractive, especially to fourth graders. However, many in both grades did not have a clear idea of the effects of weights placed

on either side. This apparent lack of familiarity in using instruments was not limited to the balance. In Playground and Math Town the majority of students ignored other measuring devices in favor of the ruler, however inappropriate it was for the task at hand. On the other hand, students did show prior experience with bulbs and batteries and handled these pieces of equipment with ease.

Students generally enjoyed the tasks and worked well with one another. It was obvious, however, that solving problems such as these represented new experiences. Administrators noted that, when students did have prior experience with this kind of activity, it was evident in their approaches to the task at hand.

Implications for Teachers

There are two main purposes of the reports on these performance tests. One is to supply detailed information on how students understand mathematical and scientific concepts and how well they are able to apply those concepts to practical problems. We hope that this information will prove useful when teachers plan instructional activities for their students. The second purpose is to model the kind of evaluation tasks that can, and should, be taking place in classrooms. For these reasons, these reports are specifically addressed to the principals and classroom teachers who monitor the progress of their students.

Most teachers depend upon observation for the informal assessment of their students. They have, however, voiced concern about the validity of such methods in more formal evaluations. Although they are aware that judgements based solely on simple tests of facts and procedures do not give an accurate account of what their students can and cannot do, they worry about the objectivity of tests that cannot be easily marked and computed. Focussed observation of student performance can prove to be a valid and reliable assessment technique, provided the criteria are clearly delineated and fully documented.

For each of the tasks reported here administrators used a detailed scoring guide, supplemented by their own observations and records of students' conversations and strategies. Examples in our reports show how these strategies can be used to produce objective evaluations. The extent of detail required is, of course, dictated by the assessment. For informal assessment at the younger grades, short notes and observations should remain the norm. However, for more formal purposes such as grading, more detailed and explicit criteria should be used.

Secondly, in order to understand the reasons behind students behavior, it is important that students be asked to explain their reasons or strategies. Communication is an important component to this kind of evaluation. Much of students' reasoning is inaccessible to us unless we directly question them. Asking why they think as they do, in contrast to asking for the correct solution, is an important part of assessment.

Thirdly, all the tasks were developed to be completed by pairs of students and, had it been practical, more would have been included. Partially, this is because problem solving in practical situations is usually a social process. Individual problem solving is more apt to occur in school contexts than in real life. Partially, this is because students reveal more when they work together. Not only do they discuss strategies and rationale more freely with one another than they would with a teacher, they also accept and extend the ideas of others. More can be learned about students' understanding from their work with others than from their independent efforts.

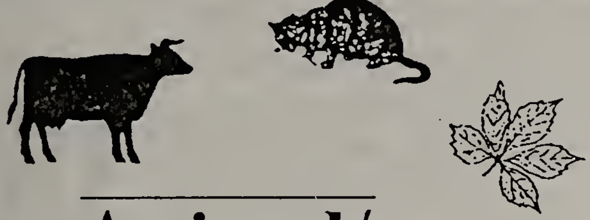
Clearly, teachers must use many kinds of information to make assessments of their students. Assessments are not tests; they are judgements about quality. To be valid they should be based on different kinds of evidence. What kind of evidence that we accept should be determined by what we want to know. For example, if we want to know how well a student can read graphs or can interpret information, then clearly a written test is the most appropriate measure. Sometimes, however, more complex tasks are needed. If we want to know how well students can structure an investigation and or can integrate different concepts in the solution of a problem—particularly a practical problem—then more active and open-ended tasks are needed. These reports suggest a beginning.

designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Circuits," "Cubes," "Insulation," "Math Town," "Playground," and "Popcorn Estimation."

BEYOND Paper & Pencil



Animal/ Leafkey

Introduction

Classification is a basic tool of science. It enables scientists to observe more precisely and to order their observations so that the relationships may be more easily understood.

We could hardly function without the ability to categorize the myriad of objects in the world around us. We learn to do this from birth, as a way of survival. As we grow older, classification helps us to store and use information. By sorting and grouping observations in different ways, we solve problems and make decisions. An ability to observe, discriminate and classify is a fundamental skill in mathematics and science. However, classification is more than ordering observations. It is also a process of naming. Without precise, well-defined terms, there are no concepts to classify.

Animal/Leaf Key, the most structured task in this series, focused on classification. It investigated students' understanding of one class of living things, the animals, and students' ability to classify different types of objects.

Problem One: Is This An Animal?

Description

Students were asked to say whether or not the object shown on each card was an animal. If there was disagreement between the pair on any card, the administrator put the picture aside and presented it again after completing the other cards. Once a decision has been made on all the cards, the administrator again presented each of the cards that children had decided were not animals and asked, "If this isn't an animal, what is it?"

Materials: Animal cards - 16 black and white pictures of the following (the name of each picture was written on each card):

car, fire, duck, tiger, snail, tree, fish, elephant, cat, worm, cow, boy, grass, spider, whale and bumblebee

A Car

Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education January 1990

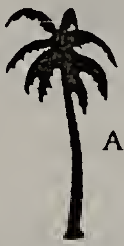
In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

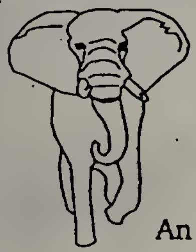
Focus

A Tiger

Many common words are used differently in science than in everyday life. **Animal** is one of them. From an early age, children form a concept of what an animal is. They are able to classify correctly animals that they have never seen, and many use the word as a generic term for anything that moves under its own power. As they get older, they learn more technical and specific names for animals, such as insects and arachnids. Later, they are introduced to the word **animal** as a scientific class. This increasing sophistication in classification and the use of language can pose problems for students and is typical of the problems posed by many scientific concepts. In science, words are used to denote something different from their common meaning, often without sufficient discussion. For example, when a person or a worm is classified as an animal, scientific terminology is likely to clash with students' intuitive ideas about what is correct. In its turn, this discord is apt to interfere with students' ability to understand the scientific concepts involved.

A Fish


A Tree


An Elephant

Results

Between 98 and 100 percent of both fourth and eighth grade students were sure that an elephant, a tiger, a cat and a duck were indeed animals. However, students at both grades became hesitant when they were shown pictures other than those of land animals. While, 85 percent of the fourth grade and 91 percent of the eighth grade realized the whale was an animal, the numbers dropped when students looked at the bumblebee. Only 64 percent of the fourth grade and 75 percent of the eighth grade agreed it was an animal. The picture of the worm brought similar results: 69 percent of the fourth grade and 73 percent of the eighth grade felt it was an animal. The picture that brought the most confusion was the boy, 26 percent of the fourth grade and 55 percent of the eighth grade agreed it was an animal. Other students used the following terms: human, man, male, person.

A Cat

A Worm

A Boy

Percentage of students replying that object was an animal.

	Grade 4	Grade 8
spider	70	75
fish	84	86
frog	87	89
whale	85	91
worm	69	73
snail	76	80
boy	26	55
bumblebee	64	75

Grass

When students replied that an object was not an animal, the administrator categorized their answers to the question **"what is it"** according to the following codes: 1. Synonym (e.g., car=auto); 2. Class (e.g., car=vehicle); 3. Specific (e.g., car=jeep); 4. Description (e.g., car=inanimate); 5. Vague; 6. Other.

Most of the children replied in terms of class (code 2) and there was a slight growth from the fourth to the eighth grade in this type of response. For example, 55 percent of the fourth grade and 60 percent of the eighth grade students gave a response such as **plant** to the picture of a tree. Others gave a specific instance such as **palm tree.** The same pattern followed when categorizing grass. Fifty-seven percent of the fourth graders and 65 percent of the eighth graders answered in terms of class, classifying it as a **plant** or using some other general term.

A Whale

On the other hand, the objects themselves affected how students responded. Fire proved difficult for children to classify. Most students used descriptive terms such as **flaming** or **red,**

A Duck

A Snail

yellow, blue when asked "what is fire?" Others described a specific attribute of fire such as **combustible.** Still others referred to its effect (**smoke, hot, dangerous**) or gave a specific example, such as **camp fire.** It seems students had difficulty finding the "right" word to clearly state their intended meaning. This may be explained by the students' lack of vocabulary or simply lack of understanding the concept of fire.

A Spider

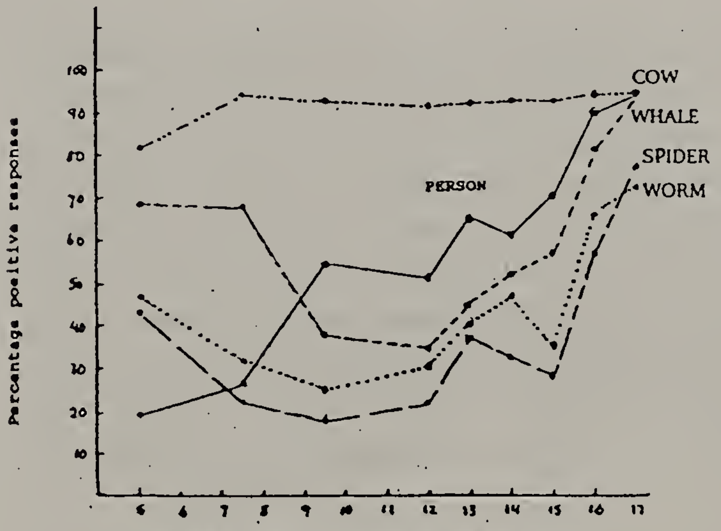
It's interesting to note that although the students did not recognize certain objects as animals, in most cases they used more specific terminology. The majority of students who did not believe that the spider and worm were animals classified them as insects. For example, 19 percent of eighth graders and 15 percent of fourth graders gave a class categorization such as **insect** or **invertebrate** to the picture of a worm. Some of those who didn't agree that a whale was an animal labelled it as a mammal. This, undoubtedly, shows the difficulty faced by students who learn more specific terminology without understanding how it relates to their existing conceptual scheme. The discussion of this 15-year old student, which is quoted by Osborne and Freyberg,[1] may express the frustrations of many students, including some in our sample.

A Cow

> ...before we started on science, when we were in the younger school, when they said animals, you sort of thought (of) all animals. You didn't sort of class insects as insects. Everything was just an animal. It wasn't human. But now...they take things away. That group is insects..and mammals or whatever...and at this age you are told that everything is just an animal or a plant...they're wasting time where they say that they are splitting everything up. And then they are just putting everything back together at the end again.

Surveys which have investigated students' conception of an animal have illustrated how understanding of what should be included in the class of things labelled as "animals" changes with schooling. The graph below shows the proportion of responses to the question "Is it an animal" for representative samples of 5 to 15 year old students as well as 16 and 17 year old biology students. In comparison, the percentage of correct responses among the students in the assessment group were higher, perhaps a partner was involved. The involvement and discussion that occurs with a partner may have resulted in more thoughtful consideration on the part of these students.

A Bumblebee



representative sample response to "is it an animal?"

The responses above indicate that children tend to be very inclusive in their initial categorizations. Then, as they learn more specific terminology, they become more exclusive. Finally, they begin to learn that the scientific definition of an animal may differ considerably from their intuitive understanding. They begin to make the distinction that, while a lion is an animal, *technically speaking* so is a worm and so is a person. The meaning of words depends on their context. It is important that they classify using both general and specific frameworks. Students need to generalize and further reclassify. They need to understand that a whale is a mammal, which is a specific type of animal; a kangaroo is an animal, yet is also a particular type of mammal called a marsupial.

Fire

---

[1] Much of the background material for our discussion on the concept of an animal was taken from Osborne, R. and P. Freyberg. **Learning in Science.** Auckland, New Zealand: Heinemann. 1985. Suggestions for student activities in observation and classification can be found in Schools Council. **Like and unlike.** London: Macdonald Educational. 1973. Other activities were contributed by Susan Angell (North Andover Public Schools), Kristen Eastman (Provincetown Public Schools), and Bernadette Golden (Winchester Public Schools).

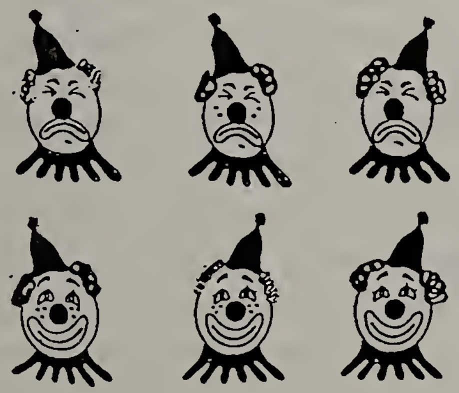
Problem Two: Classification of Clown Cards

Description

The students were asked to divide the pictures into two groups according to some attribute, and then to redivide each group. After each division, they were asked to explain the basis for their categorization.

Materials: 12 colored pictures of clown faces with the following attributes:

hat/ no hat; hat tilt right/left; smile/frown; light/dark hair; freckles/ no freckles; eyes open/ eyes closed.



Focus

The set of clowns, with their 6 different attributes, "set the scene" for the more complicated categorization of natural objects. The carefully contrived differences that they display were easily recognized and helped children to think about the logical processes involved in sorting.

Results

During the first separation of the clowns, the majority of students (62 percent in both grades) separated the clowns according to the presence or absence of a hat. Others separated according to whether or not the clown showed a smile (16% of the fourth graders, 14% of the eighth graders) and freckles (14% of the fourth and 17% of the eighth grade).

Almost all students separated clowns into mutually exclusive categories (i.e., those that represented the presence or absence of some attribute). As they continued to separate the clowns into different categories a second time the percentage dropped slightly—92% of the fourth grade students separated the clowns into mutually exclusive categories and 95% of the eighth grade students did so.

This task is typical of many "concept" type tasks that involve the conjunction and disjunction of clearly defined attributes. For example, given a small red triangular block and a small yellow square, children are often asked to recognize what attributes the blocks hold in common (size) and what differs between them (color and shape). For students who had been exposed to this type of classification activity, **Clowns** proved to be very simple.

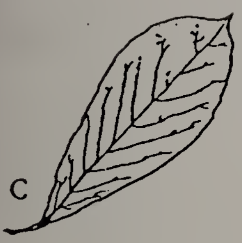
Problem Three: Classification-Leaves

Description

A

The students were asked to divide 11 leaves into two groups based on some characteristic which they observed. When both students were satisfied, they were then asked to name the characteristic which described each group. The administrator next asked them to separate each of these groups into two other sets and to describe each of these new sets. Finally, the administrator presented the 12th leaf (one that had been arbitrarily chosen from the original pack) and asked the students to place it within their classification scheme.

B

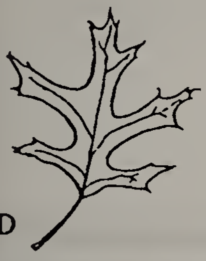

C

Students were then asked to repeat the task using different classifications.

Materials:

12 leaves, laminated in plastic with a letter attached to the front indicating its type (for the administrator's use). The leaves displayed above and below were used.

Focus

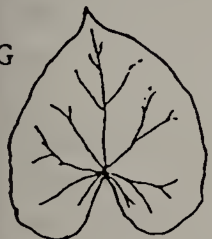

D

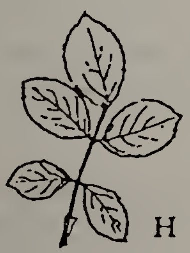
E

F

Whereas the clowns presented a clearly defined set of attributes, most of the real-world objects presented for categorization are fuzzy and ill-defined. Most classification systems proceed **ad hoc**, dependent upon the logical clarity of the creator, as well as her/his ability to define precisely the differences that distinguish one instance from another. This open-ended task was designed to investigate how clearly students were able to make distinctions between different groups of leaves and how well they were able to articulate those distinctions. It is typical of the kinds of classification problems that botanists, biologists and other scientists face when they attempt to classify real objects or events. Unlike the previous task which involved the presence or absence of a set of clearly defined attributes, some leaf attributes are mutually exclusive (serrated/not serrated), while others are relative (color, size and shape) and a matter of judgement.

Results


G

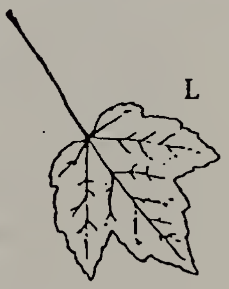

H

I

J

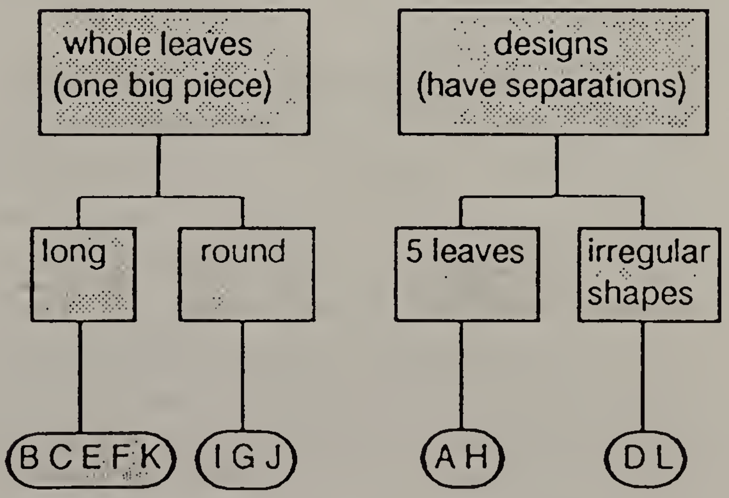
In devising a classification key, which strives for clarity, it is more efficient to concentrate initially on those attributes that are mutually exclusive and, therefore, easily recognized. For example, most plant keys start off with a division between simple and compound leaves. Only later do they use more relative attributes such as shape or size. Students seemed to sense this. Approximately half the students at both grades made their first separations in terms of the presence or absence of a clearly defined attribute (simple versus compound; serrated versus smooth edged). Twenty-one percent of the fourth and 22 percent of eighth graders divided the leaves into those that were serrated and smooth-edged. Another 25 percent of eighth grade students and 14 percent of fourth grader used compound versus simple as the distinguishing attribute in their classification. Color and size were more often used during the second divisions.

K

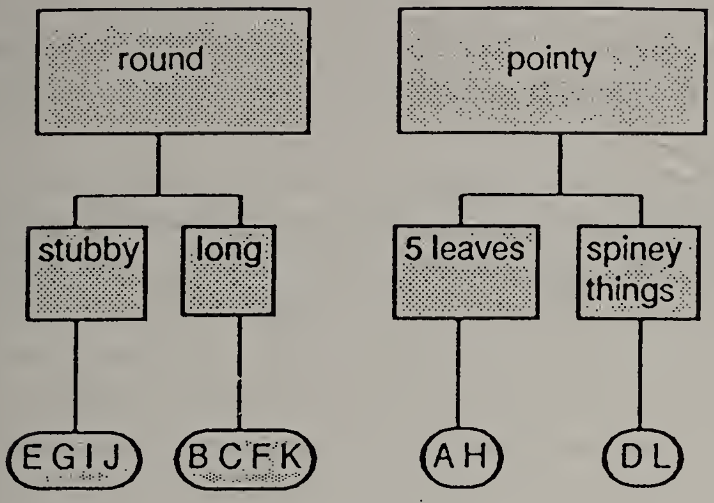

L

Whatever their initial categorizations, students tended to see the same features in the leaves and eventually placed them together in the same sets. For example, B C F and K were often placed in the same set and described as "feather-like." H and A were placed together, as were D and L. Students' chief difficulty lay with the language. They did not have the vocabulary to describe accurately many of the attributes that they observed. Although fourth graders' descriptions of their categories tended to be more general than eighth graders, many at both levels experienced difficulty in expressing the rationale for their separations.

Two typical examples are as follows. The first is the classification of a fourth grade pair of students.



The second is the classification of two eighth graders.



Attitudes

Overall, students seemed to enjoy these tasks, feeling they were "fun." For most students there was a high degree of involvement at both grade levels. (Approximately 89% of fourth graders and 84% of eighth graders remained engaged throughout completion). However, administrators noted that students at grade four tended to be more enthusiastic about the tasks than at grade eight (85 percent versus 73 percent). This may have reflected the fact that some of the older students became bored with the second repetition of the Clowns and Leaf Key and were, consequently, less accurate or concerned with accuracy the second time.

A common theme among administrators' comments concerned students' tendency to act first, think later. This was especially apparent in the Leaf Sorting task, where students established their groups with minimal consultation with one another and, apparently, little thought. In contrast, the question of whether an object was or was not an animal led to interested discussion among some children. One administrator noted:

> They debated the categories for those things which were not animals. They felt embarrassed at first, but really got into it and seemed to enjoy explaining to each other *why* their choice of classification made sense. Much learning seemed to be taking place in the debate...even to the sharing of definitions and ranking of the levels of categories.

The tasks raised other questions as well. Not only were some students interested in learning more about the leaves (wanting to know what type of trees the leaves were from), others were interested in knowing if previously tested students had given "better" categories than they.

Summary

In contrast to many of the performance tasks, **Animal/Leaf Key** examined students' language and concepts, as well as their ability to solve problems. Although it appeared to be a very simple, rather unchallenging task, it yielded valuable information about students' difficulties in understanding scientific terminology that adults often take for granted. It also revealed the differences between different types of classification tasks.

Language

Students are continually being introduced to new terminology in their science instruction. It has been calculated that, during the years from the kindergarten to twelfth grade, students are presented with more new vocabulary in their science curriculum than in their foreign language instruction. Add to this the additional complication that scientists often use familiar words in special ways. It is no wonder that students are often hazy about the meaning of terms. Furthermore, students' misunderstanding of scientific terminology is not always evident in the ordinary classroom. When terms have different meanings for students and teachers, it is not always apparent. It *is* important, however. When people hold different meanings for the same word, communication is no more possible than if they are speaking different languages.

Animal is a good example of the difficulties that can occur. Children begin with a very general inclusive meaning to the word: animal can be anything on four legs. They then are taught specific words for different types of animal. Finally, they learn that scientists' idea of animal is different from what most people would call an animal. How confusing! Does one term supercede the other? Or, are there different ways of

answering in class and in everyday life? Although some students seemed to understand the classification systems involved, many did not. This illustrates the need to discuss and make explicit how different terms are related. Because language is so natural, this may not be appreciated by those who are already familiar with the terminology.

Language problems, if not terminology, were apparent in Leaf Key. Students sorted their groups but found it hard to define the criteria for their classifications. They lacked the vocabulary to describe the characteristics of the leaves with any precision. This leads us to question how much they actually observed about the leaves. Precise and specific terms help us to focus our attention on a specific aspect of what we are observing and help us to observe better. In fact, phenomenon becomes more interesting only when we observe it closely and are able to describe it accurately. Again, it seems that the skill of observation is also related to language and is fostered in conversation, discussion, and trying to communicate with precision and accuracy.

Classification

The potential problems in classification are not trivial. The best systems are not obvious, as witnessed by the number of existing plant identification keys, each with its own set of advocates. Students should understand that classification systems are arbitrary and dependent upon the use to which they are put. For example, animals may be classified according to vertebrates and non-vertebrates, and then into birds, mammals, reptiles, etc. However, there other equally legitimate ways of classifying animals, such as the type of place where they live. What is constant among classification systems is the progression from general to particular and the requirement that any specific level must include all instances contained at that level. It is also important that students understand the hierarchical nature of classification, and that the concepts used in the initial divisions are usually the most basic distinctions that can be made within the group. For example, to sort the leaves in terms of size, a relative characteristic, is less effective than to sort them according to their basic shape (i.e., compound or simple). As in the case of scientific terminology, the conceptual nature of classification demands a lot of discussion as well as observation.

Extension for Classroom Instruction

Students need time to analyze, evaluate and discuss. They need to develop a listening appreciation of the viewpoints of others. Students also need time to discuss "why." Through classroom discussion student gain a greater appreciation for new ways of thinking. This is a valuable learning process as they are made aware of how and why other viewpoints may differ from their own. By discussing the reasons "why" the student decides an animal is or is not an animal, the students (and teachers) are open to new ways of thinking.

Students also need practice in the process of observing, defining and ordering the distinctive and common properties of different objects. Listed below are some activities that involve observation, sorting and grouping within the context of scientific exploration. Although most are appropriate for the earlier grades, all might be thought of as starting points for older students.

Grouping. Students can collect other things besides leaves—buttons, animal pictures, leaves, screws, pieces of cloth—and sort them into sets according to some contrasting properties. What are the different ways that these objects can be classified?. Compare and discuss. Are there any "right" or "wrong" ways to classify? Are some ways better than others? How many different ways can the same group of objects be classified?

Contrasting Groups. Let children choose their own contrasting properties and ask them to collect objects that exemplify those properties. For example, smooth/rough; things that float/things that sink; things that bend/ things that are rigid. Can all the objects in each collection be exclusively categorized according to one or the other property? For example, does a paper boat always float? Are there some rocks that are hard to classify as either smooth or rigid? Why is it easier to classify according to some properties than according to others? This may lead to a discussion of the difference between absolute and relative attributes.

Overlapping Sets. Have children make up sets of things which have something in common: red things, woolen things, bottles, stamps. How do they sort these out? Where are the overlapping sets? Drawing Venn diagrams helps here.

Odd Man Out. Give the students a set of buttons and have each isolate one that is slightly distinct from the others. Who can identify the one that doesn't belong?

Ordered Sequences. Ask children to classify things according to some relative dimensions such as size, weight, volume, shading. Can we order objects along a continuum? This calls for some practical measuring activity. How can we tell which leaf is larger? How can we measure the density of wood? What about the volume of different containers? What about things that are not regularly shaped?

Making Discriminations. The ability to make fine discriminations in order to classify objects or occurrences underlies most scientific experimentation. Ask students to sort and classify similar objects that show fine distinctions: smooth stones, glass bottles, cans, screws. Next ask them to do the same with objects that appear identical but, on closer inspection, show interesting variation: leaves from the same plant; seeds from the same packet. The use of a magnifying glass or a X15 or X20 microscope will help here.

Extending beyond the Class. Have students investigate how people classify things in other contexts: library books; plants; shapes; sailing boats. Catalogues can be useful here. Ask the students to identify or make up some keys for the different schemes.

Adaptation for Evaluation

Tasks such as Animal/Leaf Key set the scene for discussion and communication among children. They allow the teacher to listen and record the precision of students' vocabulary as well as the accuracy with which they use scientific terms. There can be large differences in students' ability to define terminology and their ability to use it correctly. By listening and recording how students express their ideas, teachers can learn more about students' scientific understanding than they can by using a short-answer written test.

An example of an evaluation task is given below. Have students:

1. List the major characteristics of plants and animals.
2. List all the plants and animals that are known.
3. Test each characteristic to see if it makes a clear division between all the listed plants and all listed animals.
4. Choose plants or animals. List known or observed characteristics of the set. Test the generality of the characteristics by attempting to make a division that will separate all the listed instances into two distinct groups.
5. Continue with each of the subdivisions, first listing characteristics before testing their usefulness by actually placing the instances.

Note the number of characteristics students list, the accuracy and precision of their observations, how well they recognize general defining characteristics, their ability to foresee or recognize overlapping categories at lower divisions, and their ability to recognize redundancies. Some students may know the terminology associated with the traditional subdivisions; others may not. However, it is important to stress that this is not a test of prior knowledge about classifications systems, that things may be classified in different ways as long as the system is hierarchical and that all instances are included (and not duplicated).

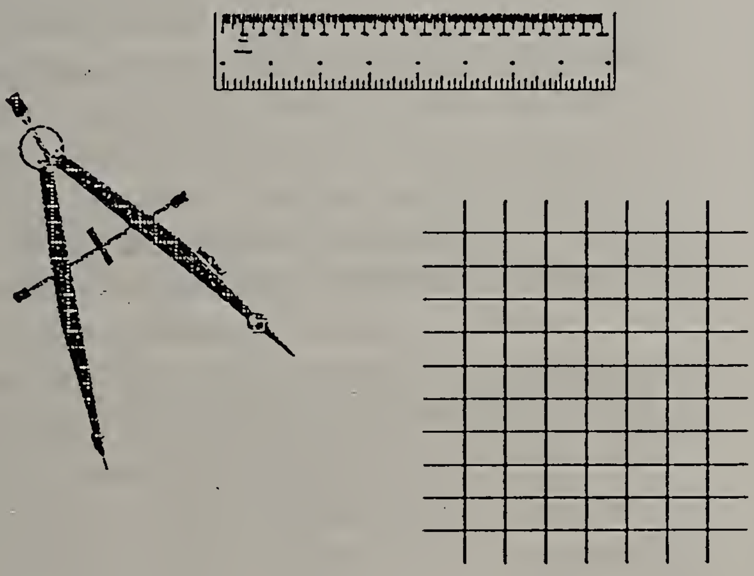
designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Background Summary," "Circuits," "Cubes," "Insulation," "Math Town," "Playground," and "Popcorn Estimation."

BEYOND Paper & Pencil

Playground



Introduction

Playground was a companion to the eighth grade task Math Town without its complexities. Still, this task required a transposition from the literal to the figurative in planning the playground. Using a map of a playground allowed students to work with something familiar, making it easier for them to visualize the whole.

Focus

Playground examined fourth grade students' abilities to organize and deal with logical and practical constraints, while looking at such mathematical concepts as area, perimeter, locus and scale. It also investigated students' skills in computation, measurement (accuracy and precision), and the correct or incorrect use of instruments.

Description

This task was divided into two parts, the first asking students to design a playground. They were given a base map of a proposed playground where they were to position equipment pieces according to certain restrictions. They also were given pictures of the following playground equipment: a general open area, basketball court, sandbox, swing, monkey bars and slide. Students were asked to place the equipment in the best spot, keeping the restrictions in mind.

Materials

- Playground map
- Ruler
- String
- Compass
- Transparent grid

Massachusetts Educational Assessment Program

Massachusetts Department of Education January 1990

In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by

Administrators read aloud the following restrictions in order to alleviate any problems students might have with reading. Students were also given cards with these restrictions printed on them:

> The slide and the swings must be at least 1 inch from the open space, the basketball court, the jogging path, the fence and the other playground equipment.
>
> The monkey bars and the sandbox must be at least 1 inch from the jogging path.
>
> The basketball basket must be at least 1 inch from the jogging path and other equipment.

The second part of the task consisted of an area and perimeter determination exercise. Students were asked to answer the following short questions.

> What is the area of the open space in square inches?
>
> How long is the jogging path in inches?
>
> If a jogger starts at entrance A and jogs 105 feet, where on the map will she stop if the scale of the map is 1 inch = 5 feet?

Results

Placement of Playground Equipment

In placing the playground equipment, all of the tools provided could be used, but some were better than others for meeting the different requirements of the various equipment. Administrators noted which of the tools students used.

The initial strategy used by 56% of fourth graders was to place the large pieces of playground equipment first.

Open Space

Since there were no restrictions on it, students were able to position the open space with few problems. The open space was used as a restriction for other equipment.

The basketball basket must be at least 1 inch from the jogging path and other equipment.

About half the students (52 percent) understood that the restriction in this question applied only to the basket, not to the entire court, and since it was an unconventional court with one basket, students only had to make sure that the basket was one inch from the jogging path and from the other equipment. The ruler or the grid was the best tool in placing the basketball court, since only a simple measurement was needed. Of the 64 percent of the students who used the ruler to place the basketball court, 66 percent placed it correctly. Seven percent of the students used the grid for this task and two thirds were successful in their placement. In all, slightly less than half the students (48 percent) were able to place the basketball court on the playground correctly.

The monkey bars and the sandbox must be at least 1 inch from the jogging path.

Although the grid or compass were the best choices for placing these items, most students used the ruler. In locating the sandbox, 63 percent of the students used the ruler and 56 percent of them were correct. Sixty-one percent of the students used the ruler to place the monkey bars and 65 percent of them placed it correctly. Overall, 38 percent of the students correctly placed the sandbox and 49 percent placed the monkey bars correctly. It's difficult to account for the different success rates here, except to comment that students were generally more successful in placing the monkey bars regardless of the tools used. Size may have been a factor. It may have been easier to find an area that accommodated the smaller monkey bars than the sandbox.

The slide and the swings must be at least 1 inch from the open space, the basketball court, the jogging path, the fence and the other playground equipment.

These two pieces of equipment had the most restrictions on where they could be placed since essentially they had to be an inch from everything else. The compass would have been the most efficient and accurate way of placing them, but only one percent used it for either of the placements. About 60 percent of the students used the ruler with 65 percent of them placing the swings accurately and 54 percent placing the slide accurately. In all, one third of the students successfully placed the swings and just over one third (35 percent) located the slide correctly.

Use of Tools

Students were hesitant to use all of the tools for these placement tasks. They tended to use what was most familiar to them, the ruler, whether or not it was appropriate. Sometimes even familiarity with the tools did not help their cause, as in the two examples reported below.

> Both boys seemed comfortable using a ruler, yet they were inconsistent with their results. I believe they were hampered by the fact that they did not seem to recognize the value of keeping the ruler flat while taking their readings. Also, they seemed able to focus on one restriction at a time but not several simultaneously.
>
> The selected student noticed the ruler use would be easier if started on the one inch instead of on the 12 inch end. One of the students asked 4 1/2 or 4 which is it? "4" - They didn't have the concept of 4 1/4. Confusion occurred when one measured and one recorded. Measurer doubted partner was recording correctly.

Approximately 60 percent of the students used the ruler for everything. The following chart lists the tasks and the tools used by the students.

	Ruler	Compass	String	Grid
	%	%	%	%
Open Space	61	1	1	7
Basketball Court	64	1	1	5
Sandbox	63	2	1	5
Swing	61	1	1	7
Monkey bars	61	2	1	5
Slide	61	2	1	5

In the following example, the two students had no trouble understanding and abiding by the restrictions given. However, they did not measure accurately; they used estimation as their means of measuring rather than using the instruments provided for the different placements.

> 1. Placed basketball court and measured (biggest things first).
> 2. Placed open space; discussed location of net vs. path.
> 3. Placed swings, and measured distance from jogging path.
> 4. Removed all but swings (put smalls in first).
> 5. Placed slide, moved swings, placed monkey bars (no measurement).
> 6. Placed open space, tried to place basketball, removed everything.
> 7. Placed open space, basketball court, swings, slide and monkey bars.
> 8. Placed sandbox.
> 9. All the while no measurements (all by eye),
> 10. Checked swings measurement from basketball court.
> 11. Moved swings a little; "that's better."

Overall, students seemed to like the idea that it was a playground they were designing. In some cases the students were more concerned with the actual playground designing than meeting the given restrictions. One administrator related the following example: "As time passed the partner wanted to color on the map or add other items like benches - this was distracting to the selected student...They seem disinterested at this point." Other students were quite concerned with the practicality of the park.

Area and Perimeter

What is the area of the open space?

About half the students were able to answer this question correctly. Of those students who correctly responded, 81 percent used the grid and counted squares and 17 percent measured and calculated correctly. Students seemed quite willing to use the grid in this context. Almost half the students (47 percent) used the grid and 86 percent of them were able to find the correct answer. About one third of the students (31 percent) attempted to measure and calculate the area. Students who used this method were far less successful with only 27 percent getting it correct. Most of the students who tried this method failed because of their calculations, not because of faulty measurements.

How long is the jogging path in inches?

Students used various methods to find the answer to this question. Almost 70 percent used the ruler to measure each side and added the lengths of the sides; 8 percent used the ruler to measure two congruent sides, multiplied those sides by two, and added the totals; 14 percent used the string to trace the outline of the path and then measured the string; and 3% used the grid and counted the squares. About half the students (48 percent) gave the correct answer. It appears that there is no one strategy better than the others to find the answer to this problem. In all the strategies used, students generated the correct answer more than 40 percent of the time. Students were most successful when they measured each side, 51 percent of the students who tried this method succeeded.

The administrator read the next question only after students had answered the previous question.

> If a jogger starts at entrance A and jogs 105 feet, where on the map will she stop if the scale of the map is one inch equals 5 feet?

This was a more difficult question not only because was it a two part question, but also it required proportional reasoning. Slightly less than half the students (46 percent) were able to answer it correctly. Of those students who did respond correctly, about 70 percent counted inches by fives and stopped at 105. Another 20 percent answered correctly by dividing 105 by five and measuring the correct number of inches. The remaining ten percent used other methods most often characterized as "accidentally correct." Overall, 57 percent of the students counted by fives to find where the jogger would stop and 57 percent of them gave the correct answer. Far fewer students (ten percent) divided 105 by 5; however this technique resulted in a higher percentage of correct answers with 88 percent of the students who used this method finding the right answer.

Attitudes

Seventy-five percent of students shared the tasks with one another while 10 percent worked independently on the same sub-task and 4 percent worked independently on different tasks. Students seemed to enjoy the challenge of this task with 82 percent of them remaining engaged until completion. About 40 percent of the students were careful and checked their results, 30 percent were careful but did not check and 23 percent showed little concern for accuracy. Sixty-six percent of the students were confident in their ability to proceed to find a solution, 23 percent displayed some uncertainty and 7 percent showed a substantial lack of confidence. Three-fourths of the students appeared enthusiastic to the test administrators.

Extension for Classroom Instruction

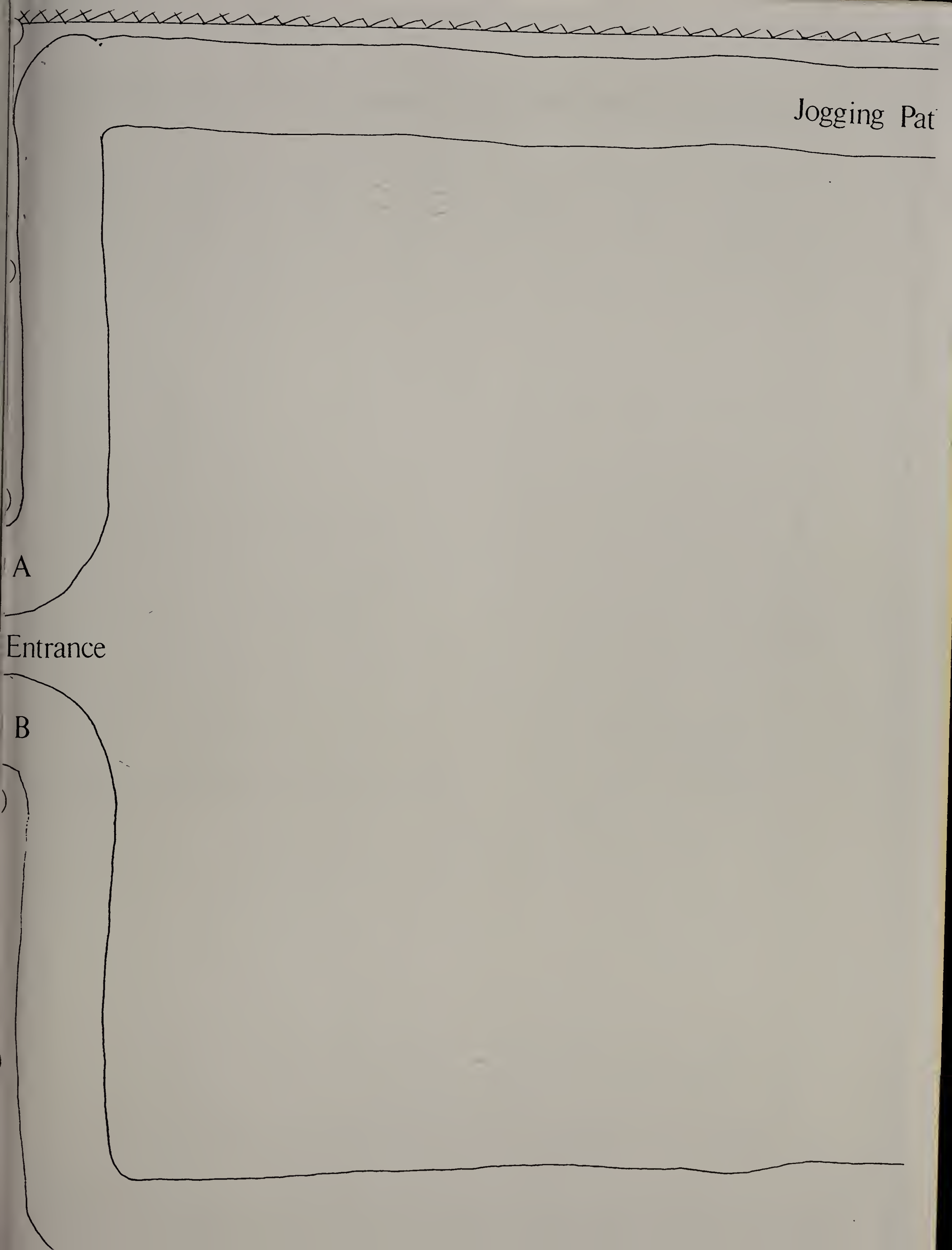
Students should begin to understand the nature of measurement as the repeated application of the same unit. By studying the nature of measurement as they learn the techniques, student will be more willing to experiment with unusual forms of measurement while understanding that measurement is more than inches, feet and yards. Students can use unconventional forms of measurements such as handlengths, coins, or pieces of paper. They could compare how many of each unit are needed to measure the same thing. The different answers resulting from different forms used might be compared and contrasted. Students then should try to make general rules about measuring.

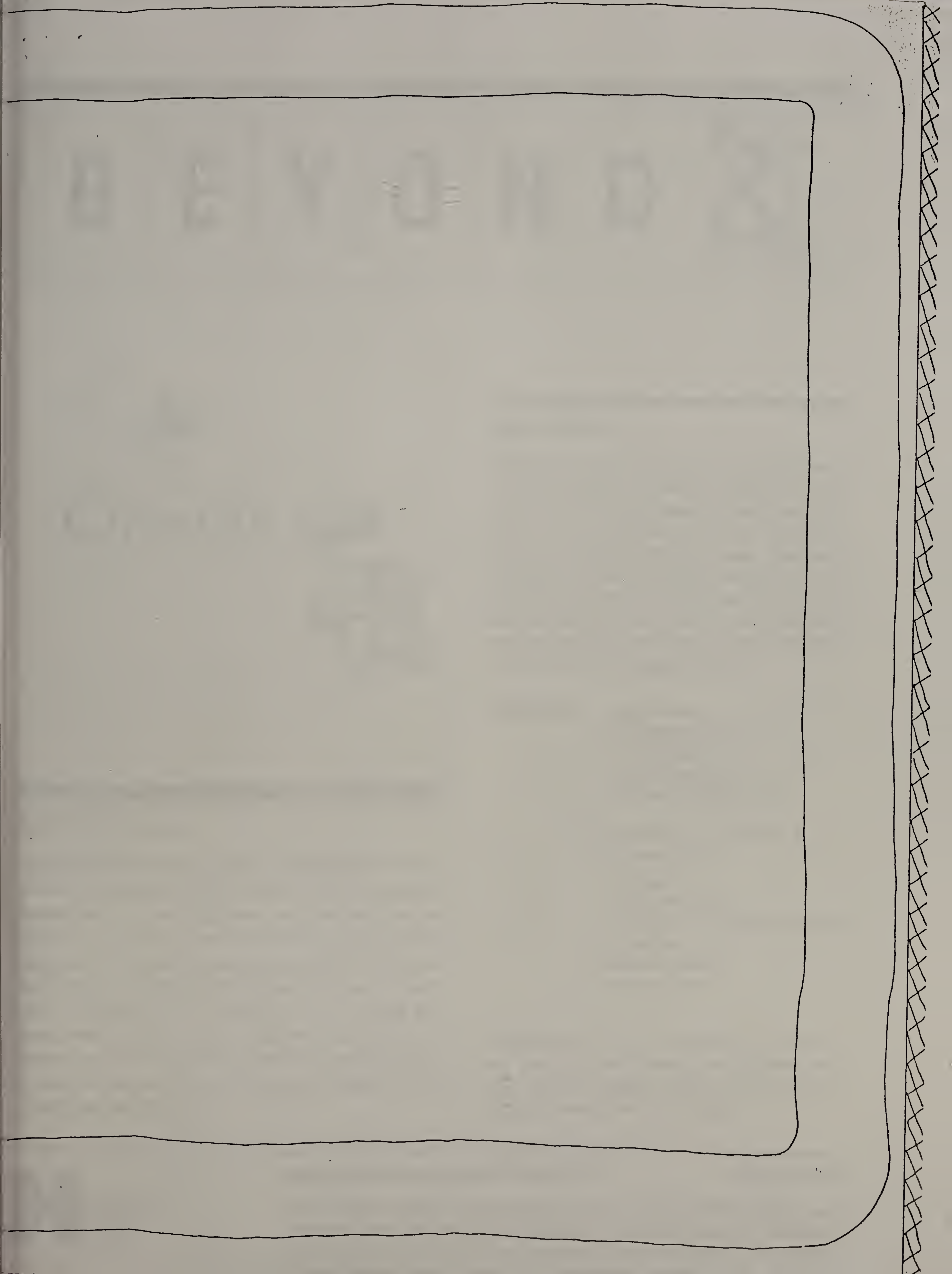
They might try measuring things that seem impractical to measure. For instance, how would they measure the perimeter of the playground without using a conventional means of measurement? They could try to relate their unconventional means to a conventional means of measurement. How do they measure around the trunk of a tree or an irregularly shaped object? They could use surprising units of measurement. For example, by making a mark on a bicycle wheel and counting the revolutions of the wheel, a new unit of measurement can be used. This would be a means of introducing students to the use of ratios and scales.

designed by

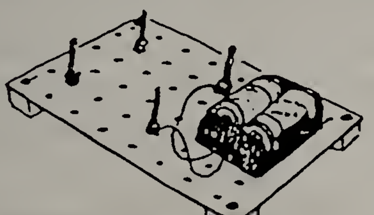
Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Background Summary," "Circuits," "Cubes," "Insulation," "Math Town" and "Popcorn Estimation."

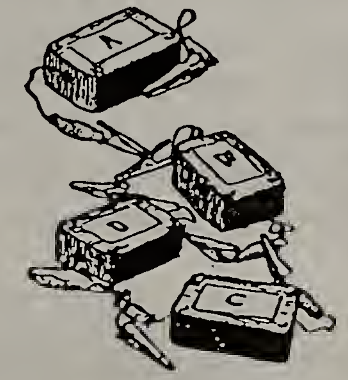

Jogging Pat
A
Entrance
B



BEYOND Paper & Pencil



Circuits



Introduction

The focus of this task was students' comprehension of the flow of electricity and simple electrical circuits. The task considered students' ability to form hypotheses and experiment with the materials in order to draw conclusions based on those hypotheses. This task also measured students' observation skills and their ability to either change their hypotheses in accordance with results, or use their hypotheses in order to give their results some structure. Plus, we observed their methods in generating responses. In this task, if students did not appear able to insert an unknown into the circuit appropriately, the test administrator demonstrated it.

Description

This task consisted of three related problems. The first problem required students to construct a circuit when given various materials such as a battery (in holder), wires, and a light bulb. The second part of the task asked students to make another circuit, this time using a partially constructed circuit board along with the same materials. In the third part of this task, students had to use the circuit board to identify the contents of four mystery boxes.

Materials:

1 circuit board
1 "C" battery holder
2 "C" batteries
1 1.5 volt mini-light
2 alligator clip connectors

4 unknown boxes labelled A, B, C, D containing:
1 AA battery
1 100 ohm resistor
1 straight wire-to-wire connection
1 toothpick
paper and pencil
unknown sheet

The unknown boxes were constructed in such a way that there were wires coming from each side. We added weights to all boxes so they were all approximately the same weight.

Massachusetts Educational Assessment Program

Massachusetts Department of Education **January 1990**

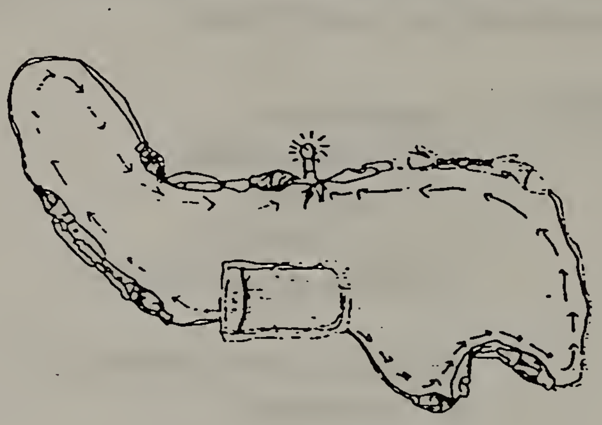
In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

Understanding Circuits

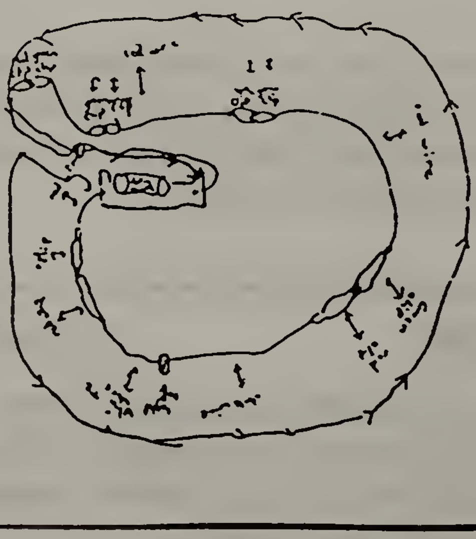
Problem 1—Results

The first task required students to light the bulb using the materials presented (battery in holder, wires, and light bulb). Students did not have to use all the materials. The administrator only interfered when it appeared that the equipment might be seriously damaged. This seldom happened. At both grades (4 and 8), students were able to complete this part of the exercise successfully; 95 percent of the fourth graders and 97 percent of the eighth graders lit the bulb. Most of them created simple series circuits (Grade 4, 94 percent; Grade 8, 98 percent). The others created parallel circuits. Only 19 percent of the fourth and 5 percent of the eighth graders completed this task by trial and error.

Students were asked to draw a picture of the circuit they had made and to use arrows to trace the flow of electricity. At Grade 4, 63 percent of the students drew arrows that converged at the bulb. Forty-five percent of the eighth graders answered in this manner. See the meticulous example drawn by the fourth grade boys below.



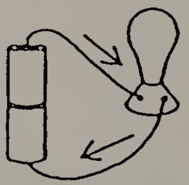
Thirty-two percent of the fourth graders and 54 percent of the eighth graders drew the arrows continuing around the circuit. Two eighth grade girls provided the following example.



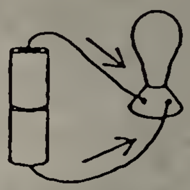
When asked how the electricity lights up the bulb, 42 percent of the fourth graders and 62 percent of the eighth graders answered that "Electricity passes through the bulb."

One fourth grade girl made the following analogy to describe how current travels in a circuit: "the current was like a car travelling down a road, when the bridge was up, the car had to stop, when the bridge was down, the car could get around and the current could pass." Student responses are nearly identical to those in a multiple choice question from the pre-performance test.

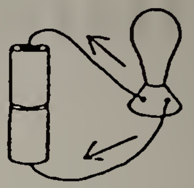
A.



B.



C.



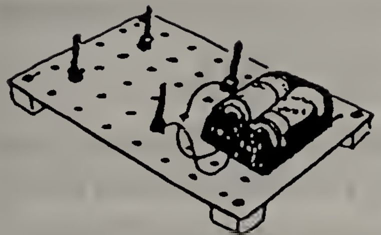
Which picture shows the direction of the electrical current in a circuit?

	Grade 4	Grade 8
A.	42%	60%
B.	40%	34%
C.	15%	5%

When asked to tell how electricity lights the bulb, 56 percent at Grade 4 and 35 percent at Grade 8 responded that "electricity stops at the bulb."

Problem 2

Next, students were given a partially constructed circuit board and asked to light the bulb using the circuit board and the equipment they had been given for the previous exercise. The circuit board appeared to make no difference in their performance. About 97 percent of the fourth graders and 98 percent of the eighth graders were able to get the bulb to light.



This time, 84 percent of the fourth graders were able to light the bulb purposefully as opposed to 77 percent in the first exercise, suggesting that they had transferred what they had learned from the first activity. When asked to show on the board, by pointing, where the electricity comes from to light the bulb, 36 percent of the fourth graders and 54 percent of the eighth graders showed electricity continuing around the circuit. 56 percent of the fourth graders and 43 percent of the eighth graders showed it stopping at the bulb. It is interesting to note that 8 percent of the eighth graders reversed their previous response and stated that the electricity stops at the bulb rather than passing through the bulb.

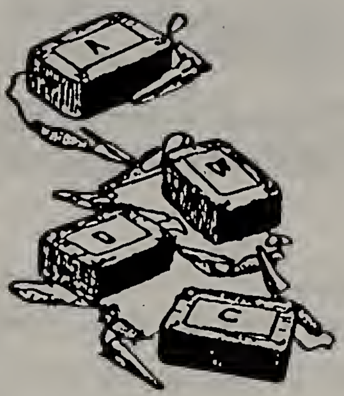
Students then were asked to try to light the bulb a second way. (There were many options here: they could add more wires, or subtract wires, or reconfigure the wires.) About 96 percent of the fourth graders and 97 percent of the eighth graders were able to do this. Most of them created series circuits by adding or subtracting wires. Only 3 percent of the fourth graders and 4 percent of the eighth graders created parallel circuits.

It is clear that although approximately half of the students did not completely understand how electricity flows, they are able to put together a circuit with little trouble. Many children at both grades argued their partners into or out of the correct answer for the circuit questions. It seemed to be counter-intuitive to children, especially those who haven't studied electricity, to believe that electricity flows in a circle and passes through the light bulb, thus causing it to light. They were more inclined to believe that it stopped at the bulb and was changed into light.

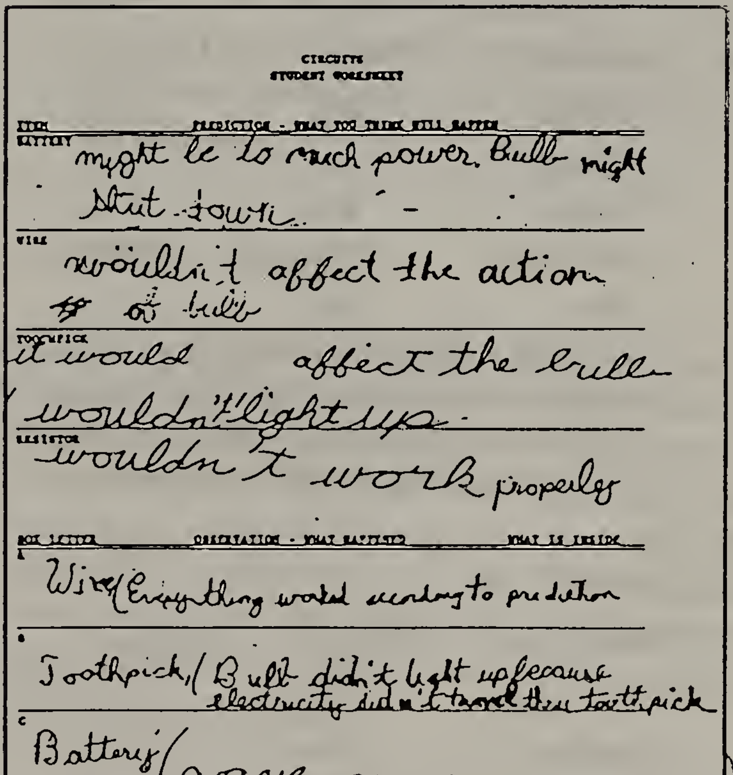
Identifying Unknown Boxes

Problem 3

The next part of this task involved formulating hypotheses, testing the hypotheses, interpreting the data, and drawing conclusions. Students had already finished three generally successful separate attempts at creating a complete circuit. This prepared them for the task at hand: that is, finding the contents of the unknown boxes.



The administrator gave the students box A (containing the wire) and asked them if they could insert it into a circuit so that whatever was in the box would affect the light bulb. Students who were unable to do this were assisted by the administrator who demonstrated the appropriate way to insert the box. The administrator, while disconnecting the circuit, said to the students that there were four things in the boxes: a battery, a resistor, a wire, and a toothpick. "They are all wired to the wires you see sticking out of the boxes. What effect do you think each would have on the light bulb if you inserted it into the circuit? Write your prediction on this sheet." (Students were given a sheet of paper, the top half of which was divided into four parts so that students could write what their predictions were.) Below is a completed example of a student prediction sheet.



CIRCUITS
STUDENT WORKSHEET

ITEM — PREDICTION - WHAT YOU THINK WILL HAPPEN

BATTERY
might be to much power. Bulb might Shut down.

WIRE
wouldn't affect the action of bulb

TOOTHPICK
it would affect the bulb wouldn't light up.

RESISTOR
wouldn't work properly

BOX LETTER — OBSERVATION - WHAT HAPPENED — WHAT IS INSIDE

A
Wire (Everything worked according to prediction

B
Toothpick, (Bulb didn't light up because electricity didn't travel thru toothpick

C
Battery (gave more power)

D
Resistor (knot much power)

Results

Students were encouraged to discuss their ideas with each other and come up with a joint prediction.

Wire

Prediction	*Grade 4*	*Grade 8*
Makes light	72 %	82 %
No light	9 %	5 %
* Other	11 %	2 %

* Generally when the student response was classified as "other," the student said that the bulb would be dimmer because it would take the electricity longer to get to the bulb with the extra wire attached.

Toothpick

Prediction	*Grade 4*	*Grade 8*
No light	72 %	85 %
Makes light	15 %	9 %
* Other	5 %	2 %

* This generally consisted of the notion that there would be smoke because the electricity would cause the toothpick to ignite.

Battery

Prediction	*Grade 4*	*Grade 8*
Brighter	20 %	22 %
Dimmer	1 %	1 %
* No light	9 %	6 %
Makes light	59 %	68 %
Other	3 %	1 %

* A small but significant number (9 percent at Grade 4 and 6 percent at Grade 8) said that there would be no light because the battery would increase the current too much, causing the light to burn out.

Resistor

Prediction	*Grade 4*	*Grade 8*
Don't know	11 %	11 %
* Brighter	1 %	2 %
Dimmer	1/3 %	18 %
No light	26 %	42 %
Makes light	27 %	15 %
Other	13 %	5 %

* Most children reasoned some definition for "resistor" by using the root word "resist"; hence, "become brighter" would not be an appropriate option.

Testing the Hypothesis

Students were next told to test the boxes to find out their contents. The first thing almost every student did was to shake the boxes and then try to judge the weight differences, not knowing that we had secured weights to each box to control for this. Of all the students who took this test at both grades, only one pair of fourth grade boys came up with a method to determine the contents of one box without testing it. According to the test administrator, "They twisted the wires and identified the box containing the wire because the wires went right through the box, and twisting one end affected the other."

Students were not obliged to use the circuit board but merely to identify the contents of the unknown boxes. Many students (25 percent at Grade 4 and 26 percent at Grade 8) were able to find the battery without using the circuit board by testing each unknown box independently with the light bulb and concluding correctly that the battery was the only one to make it light.

At both grades, the majority of students proceeded to determine the unknowns in an orderly fashion. About 81 percent at Grade 4 and 85 percent at Grade 8 tested one unknown box at a time, noted its effect on the circuit, and made their decision based on their observation. Those who did not do so generally tried to attach more than one unknown box at a time, or used an open circuit, or did not record their results properly.

These fourth grade girls were extremely careful in testing the unknown boxes.

---

1. Students were unfamiliar with the word "resistor," but felt it meant hold something back, so a resistor must hold electricity back and make light flicker.
2. They put each box into the circuit, from A to D.
3. After C made the bulb brighter, they shook the box.
4. Retested B to make sure it still didn't light the bulb and decided it had to be the toothpick.

5. Retested A and decided it was the wire. ("Wires make lamps light up, so why not here?")
6. Retested C by using just it and the bulb, and found that the bulb was still lighting up and declared that it proved that C was the battery.
7. Retested D and discussed whether it was the resistor or the wire.
8. Retested A and D several more times before deciding that D was the resistor.

The next example of fourth grade boys demonstrates a methodical yet careless approach to the problem.

The selected student proceeded very deliberately and confidently from one box to another. His friend acted as an assistant who would have the final say/comment without regard to evidence. At no time did either consider trying to isolate the battery. They just put each box into the circuit board and were not troubled when results were the same or different from predictions. After writing down observations, they ended up just guessing by trying to shake or weigh the different objects.

Drawing Conclusions

Toothpick - Box B

Students were most successful at finding the toothpick among the unknown boxes. Seventy percent of the fourth graders and 76 percent of the eighth graders identified the toothpick. The most frequently incorrect answer for Box B was resistor, an obvious choice since it would "resist the electricity." Sixteen percent at Grade 4 and seventeen percent at Grade 8 decided that it was the resistor. Four and two percent respectively thought it was the battery, using the logic that "two batteries would burn out the circuit."

Resistor - Box D

The next most successful identification was that of the resistor, with 53 percent of the fourth graders and 69 percent of the eighth graders successfully identifying it. It seems that many students found the resistor by process of elimination. Having found the battery, toothpick and wire, what else was left?

Many students had problems with the resistor because they couldn't define it. In most cases, if students were able to correctly predict what the resistor would do e.g., make the light dimmer, then they were able to correctly identify the resistor among the unknown boxes. Eighty percent of the students who correctly predicted the resistor would dim the lights identified the resistor. The next best response, particularly if students were unsure, was "I don't know." Of the students who gave this answer in their predictions, 80 percent of them correctly found the resistor.

Wire - Box A

Forty-one percent of the fourth graders and 69 percent of the eighth graders correctly found the wire. When the wire was inserted into the circuit, 71 percent and 82 percent respectively perceived that the battery made light. Others 12 and 10 percent perceive the light to be brighter, and still others (11 and 5 percent) saw something else.

Battery - Box C

Forty-seven percent of the fourth graders and 65 percent of the eighth graders correctly identified the battery in box C. About one fourth of the students at each grade level found the battery by connecting only the boxes and the bulb without using the circuit board. The only box that would light the bulb would be the battery box. The second most common response for Box C was the wire.

Overall Results

Twenty-eight percent of the fourth graders and 45 percent of the eighth graders correctly identified the contents of all the unknown boxes. At the younger level, boys were generally more successful than girls with 41 percent of the boys succeeding in identifying the unknowns and 15 percent of the girls doing so. This may be a result of experience; more young boys indicated they experiment with electricity in their play. At Grade 8, 51 percent of the boys and 37 percent of the girls correctly identified the contents of all the unknown boxes.

It appears that at both grade levels, knowledge of how a circuit works has no bearing on student performance in identifying the unknown boxes. Less than half of the students who correctly identified the unknowns had correctly described the flow of current through a circuit. At the fourth grade level, 42 percent of the boys and 24 percent of the girls who identified the unknowns were able to correctly show the flow of electricity. At Grade 8, 48 percent of the boys and 33 percent of the girls who identified the unknowns correctly could describe the way electricity travels on a circuit.

Eighty-five percent of the fourth grade boys who correctly identified the unknowns had found the battery outside of the circuit board by testing the unknown boxes with only the light bulb. Compare this with 27 percent of the fourth grade girls who correctly identified all four boxes and found the battery outside the circuit board. At Grade 8, 58 percent of the boys who correctly identified the circuits found the battery outside the board and 80 percent of the girls did so. Students were not told specifically they had to use the circuit board to find out the contents of the unknown boxes, but because the administrator demonstrated how to insert the unknown boxes into the circuit board, there was an implied specification.

Observation Skills

This task was designed to evaluate process skills as well as students' knowledge of electricity. For example, when students tested the boxes, did they differentiate between the brighter light from the box with the battery, the dimmer light from the resistor and the same light from the wire? When testing Box C, the box containing the battery, 41 percent of the fourth graders and 45 percent of the eighth graders said that the result was it made light. The battery could have either made it brighter, or made it dimmer, depending on how it was inserted into the circuit. Yet almost half at each grade did not perceive a change in brightness. When the battery, wire, and resistor were hypothesized simply "to make light," how then were students to determine which was which? Many times students said, "make light" but actually observed a degree of brightness without articulating it.

Recording and Using Predictions

We also noted whether students used the chart to help them find the unknowns. There are two separate areas to consider here. Did the students use the chart to check if the effect of the unknown boxes on the light was in accordance with their predictions? For example, if a student predicted that the light would get brighter when the battery was inserted, did the student use his/her own hypothesis to identify the unknown? When a student made predictions that were not confirmed, did the student reconsider the predictions and try again? Or did the student simply assign an identity to the unknown boxes based on some other reasoning? We asked administrators to record whether it was apparent that the students used their prediction chart to help them find out what was in the boxes. Sixty-one percent of the fourth graders and 48 percent of the eighth graders used the chart according to the administrators. That somewhat more than half at Grade 4 and slightly less than half at Grade 8 did so may not bode well for the scientific method. However, it may be that the nature of the task did not necessitate going back to the written predictions to solve the problem.

Careful checking is another aspect of the scientific method. In this particular task, checking the work would require re-inserting the unknown boxes and observing the results. Here, fourth graders fared slightly better than eighth with 54 percent checking their results in contrast to 46 percent of the eighth graders.

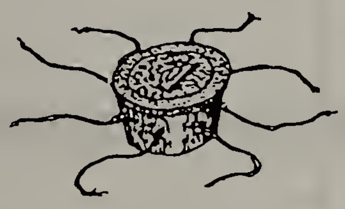
Attitudes

Generally speaking, students were at ease using the equipment. Two fourth grade girls asked whether they would "be electrocuted" if they made a mistake, but most appeared comfortable with the materials. The majority of students (83 percent at Grade 4 and 92 percent at Grade 8) worked cooperatively on the tasks. In more than half the cases (58 percent at Grade 4 and 60 percent at Grade 8) neither student dominated. Most proceeded confidently toward a solution with fourth graders (73 percent) acting slightly less tentative than eighth graders (69 percent). A large majority (91 and 93 percent respectively) remained engaged with the task until completion. Fourth graders (92 percent) were slightly more enthusiastic than eighth graders (84 percent).

Extension for Classroom Instruction

This exercise deals with the concept that a complete metal circuit is needed for current to flow.

1. Use an opaque plastic tub with a press-on lid such as a margarine tub.
2. Make eight holes (or more) around the side of the tub.
3. Thread four wires (of the same color) through two holes each. Knot the wires at the holes, either in or out of the tub, in order to prevent tugging the wire to find the other end.
4. Ask students to use the circuit board to find the ends of the same wire.

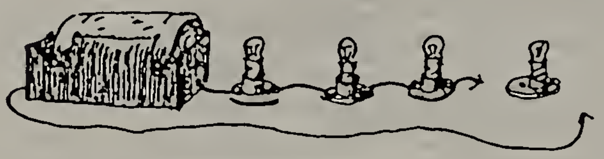


An experiment very similar to discovering the contents of the unknown boxes is described below.

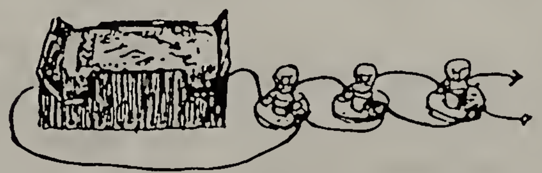
1. Set up the same circuit board, leaving a gap between two poles where students would insert the items.
2. Among items to test could be: a brass rod, a steel strip, the lead in a pencil, a piece of paper, a piece of string, linoleum, lead, aluminum, rubber tubing, plastic, plastic covered wire with and without the ends bared, tin foil, and anything else that students may want to experiment with.
3. Students should predict what they think will happen with each of the items. Try to include items that would surprise them such as pencil lead.
4. Students should keep a record of what they test and whether or not the bulb lights.
5. Then students should divide into groups those materials that conduct and those that don't conduct electricity and develop a set of characteristics for each.
6. This can lead to discussions of conductors and insulators.

This exercise deals with series and parallel circuits.

1. Give students materials such as batteries in battery holder, bulbs, and extra wires. Have students use the materials to familiarize themselves with it by lighting one bulb in a series circuit.
2. Ask students to connect two, then three, then four bulbs in a series circuit as illustrated below. Make sure they keep records of their observations of the brightness of the bulbs and the time the batteries last.
3. Ask students to find out the greatest number of bulbs that can be lit with one battery.



4. Next set up bulbs in parallel circuits. Ask students to keep records of the brightness of the bulbs and the time the batteries last.



5. Try these two experiments with batteries of different voltage. Ask students to pay particular attention to the intensity of light and see if they can draw conclusions based on the brightness.

These tasks were adapted from *Learning Through Science*, published by Macdonald & Co. Ltd.

Much of the evaluation check-list noted below is similar to that used in the state-wide administration of the test.

Observation	U/N	Comment/Example
Flow of Electricity/Current		
Continues around circuit		
Converges at bulb		
Flows in both directions on same wire		
Elementary Circuitry		
Lights bulb		
Creates a series circuit		
Creates a parallel circuit		
Creates an open circuit		
Electricity passes through bulb		
Electricity stops at bulb		
I don't know		
Formulation of Hypothesis		
Applies prior knowledge		
Considers reasonableness of theory		
Presents it clearly		
Drawing Conclusions		
Data confirms hypothesis		
Data contradicts hypothesis		
Checks data		
Adjusts hypothesis		
Arrives at appropriate conclusion		
Uses data and hypothesis		
Guesses		
Approach to Problem Solving		
Asks appropriate questions		
Evaluates different approaches		
Uses careful and methodical procedures		
Keeps accurate records		
Checks results		
Use of Equipment		
Uses with ease		
Uses properly		
Asks for assistance when needed		
Attitude		
Works with other student(s)		
Dominates other student(s)		
Shows confidence about task		
Remains engaged through completion		
Appears enthusiastic		

---

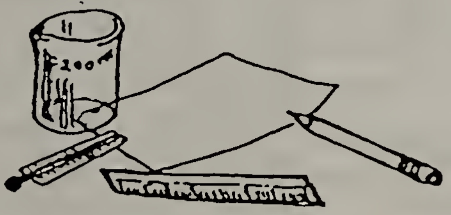
designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Background Summary," "Cubes," "Insulation," "Math Town," "Playground," and "Popcorn Estimation."

BEYOND Paper & Pencil

Insulation



Introduction

Although laboratory activities may form part of the most eighth graders' science instruction, there is little chance for experimentation. The purpose of most laboratory work is to confirm results or to provide experiences for concepts introduced previously. Seldom do students have a chance to plan and carry out a real experiment that has no clear "right" answer.

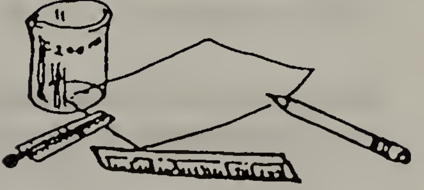
Insulation gave them this chance. It set forth an interesting question in a practical context. No special materials or laboratory equipment were involved. There was nothing to signal that this was a science test. Yet, both scientific concepts and procedures were evaluated as students investigated the relative heat loss of different hot drink containers.

Description

Students were asked to evaluate the insulating capacity of three hot drink cups and to come to a decision based on their findings. They were provided with the materials necessary to carry out and record such an investigation. The role of the administrators was limited to observing students' procedures although, at the conclusion of the experiment, they discussed the students' decisions with them.

Materials:

- an insulated plastic cup
- a plastic mug
- a styrofoam cup
- 3 thermometers
- rulers
- 1 large container filled with hot water
- a stopwatch
- plain paper
- 250 ml beakers
- a graduated cylinder



Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **January 1990**

In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

Basically, students were asked to compare the relationship between time and loss of heat for each of the three cups. In order to do this they had to identify the variables—what was to be manipulated (the independent variable), what was to be kept the same (the controlled variable) and what was to be compared (the dependent variable). They also had to consider the degree of standardization needed in order to get reliable results for the three cups.

The Controlled Variable: Quantity of Water

Seventy percent of students controlled for the quantity of water by measuring the same amount for each container. Another 7 percent appeared to recognize the need for identical quantities but used the level of the water in each container as the measurement, without accounting for the fact that each container had a different shape and, consequently, would hold different amounts of water at the same level. The rest (23 percent) did not seem to understand that the quantity of water would affect results.

In the example below, the students use a ruler to measure the amount of water in each cup. This casts doubt on their understanding of heat loss, as well as their understanding of volume and conservation. Did they notice that the cups differed in circumference? Do they realize that this difference would affect the quantity of water at different levels? If so, did they realize that a large quantity of water will retain heat much longer than a smaller quantity? Without these basic conceptual understandings, the experimental procedures themselves are meaningless.

Pours water into each cup—no measuring

Thermometers into cups

Sets timer for 1 minute

Checks temperature in each cup after 1 minute

Starts another temperature check. Puts more water in each cup after 40 seconds

Stop again (1 minute)—plastic cup worst

Styrofoam and insulated the same—checks these two. Measures amount of water in each using ruler.

Reads thermometer before putting into cups (both at 30 degrees)

Starts timer—adds more water after 50 seconds

Checks temperature after 1 minute. Insulated cup temperature is higher.

The Independent Variable: Time Elapsed

Although 67 percent of students used the same elapsed time for each of the three cups, the length chosen appeared to be arbitrary. Students did not seem to recognize that the problem was posed in a practical context, implying such considerations as: “For how long would I want the coffee to stay warm?” Had they considered such a question, they might have timed the temperature loss for more than the one or two minutes that many students appeared to believe would suffice. Only 27 percent of the students took the temperature over a time period of 4 minutes or longer.

Given the nature of the material used in the cups, this is of particular concern. It appears that the rate of heat loss for the individual cups varied over time. Whereas the insulated cup showed a drop in initial temperature as heat was absorbed by the material, it retained heat longer than the others. Only slightly more than a quarter of the students waited long enough to be able to record the different rates of heat loss for the various cups.

The students described below are more typical.

They placed the thermometer in the **styrofoam** cup and filled it. Took temperature—42 degrees, then started timer. After 1 minute it went to 41 degrees.

Then they did the **plastic** cup using the same thermometer. Poured water in, set the timer, waited a second for thermometer to adjust, took a reading at

40 degrees. After 1 minute it stayed at 40 degrees.

Did the same with the insulated cup, but it showed a reading of 41 degrees initially. After 1 minute the reading was the same.

They decided that two cups were good insulators.

The Dependent Variable: Heat Loss

Establishing the initial temperature. Because water coming from different levels in the source container could have different temperatures even if all the cups were filled at the same time, it could not be taken for granted that the initial water temperature was identical in each. Consequently, in order to establish the relative heat loss of the three cups, an initial baseline had to be established for each. Many students did not recognize this. Although 46 percent took the temperature of the water immediately after pouring it into the cups, another 42 percent took it only after some time had passed. These students merely assumed that the initial water temperature would be identical in all cups.

Charting the heat loss. Thirty-five percent of the students took only one temperature reading and 4 percent made their decisions on the basis of no readings at all. Some merely felt the outside of the cups. At the other extreme, 20 percent of the students took five or more readings in order to chart the relative heat loss of the various cups.

The students described below took no initial temperature, making their decision on the basis of a single reading. Although they used the timer, the length that they waited (20 seconds) was unrelated to the practical considerations presented in the problem. According to the administrator, they were concerned about the water cooling in the bucket, but this did not affect their decision to fill and measure each cup separately.

Measured 100 ml of water starting with styrofoam cup (felt outside of cup). Timed for 20 seconds. 45 degrees.

Measured 100 ml into plastic mug. Checked temperature after 20 seconds and felt outside of cup. 43 degrees.

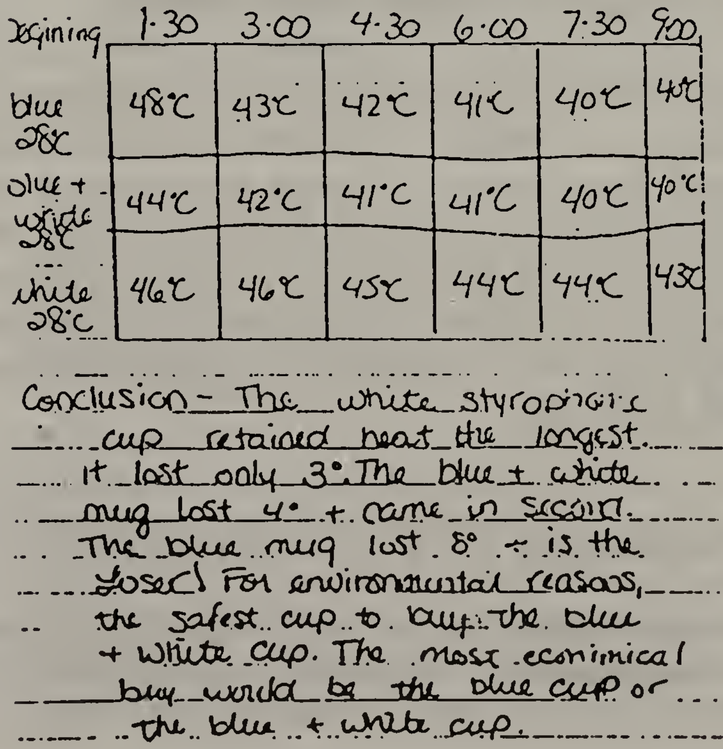
Repeated procedure with insulated cup—47 degrees.

Decided insulated cup.

Administrator's comment: They showed concern that the water was cooling in the bucket while test was going on. They also had great concern for how hot the cups felt with hot water inside. They said that the water in the styrofoam and plastic cups were losing heat to the cups and that did not happen in the insulated cup.

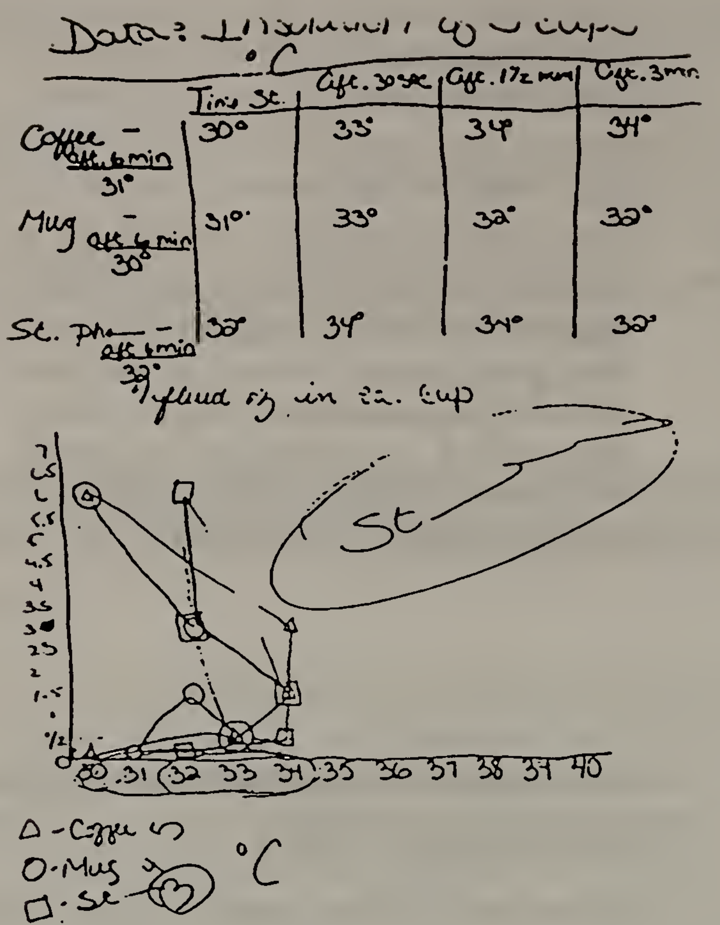
Recording of Data

The basic "stuff" of an experiment is the data collected. In order to observe a relationship it is necessary to collect and record the changes that occur when one variable is related to another. Although approximately half the students recorded some data, this consisted mainly of the final results only. Few saw the necessity of keeping track of interim data. As a result, few noted the different rate of heat loss in the different cups. The charts below illustrate the work of some students who did keep a record of their results.



begining	1·30	3·00	4·30	6·00	7·30	9:00
blue 28°C	48°C	43°C	42°C	41°C	40°C	40°C
blue + white 28°C	44°C	42°C	41°C	41°C	40°C	40°C
white 28°C	46°C	46°C	45°C	44°C	44°C	43°C

Conclusion - The white styrophoric cup retained heat the longest. It lost only 3°. The blue + white mug lost 4° + came in second. The blue mug lost 8° + is the loser. For environmental reasons, the safest cup to buy the blue + white cup. The most econimical buy would be the blue cup or the blue + white cup.


Data: Insulation of cups
°C

	Time St.	Aft. 30 sec	Aft. 1½ min	Aft. 3 min
Coffee – aft. 6 min 31°	30°	33°	34°	34°
Mug – aft. 6 min 30°	31°	33°	32°	32°
St. Pho– – aft. 6 min 32°	32°	34°	34°	32°

*fluid rz. in st. cup

Laboratory Procedures

In addition to recognizing the need for keeping the time and quantity variables constant, students had to establish a set of procedures that were precise and standardized for all the cups. Their difficulties in doing this stemmed, in part, from a lack of experience with using the instruments. For example, many did not realize that the temperature shown on the thermometer had to stabilize before an accurate reading could be made. Others did not realize that constantly taking the thermometer in and out of the water affected the accuracy of the reading.

Most students hurried through their work, intent only at arriving at a solution. The students described below are typical of many who did not seem to understand the need for precise measurement and careful recording. Not only did they assume that all the cups would have the same temperature, but they did not appear to identify that *rising* temperature was an anomaly that suggested an error in their procedure.

Measured out 150 ml, poured into insulated cup

Measured out 150 ml, poured into styrofoam. Put thermometer into insulated cup—temp = 38

Assumed temperature same in each. Allowed 2.5 minutes

Checked thermometers to make sure they were the same. Wanted to get all same (but one was touching thermometer near bulb, saying that they should be careful not to touch the bulb)

Insulated cup = 40
Styrofoam = 42
Plastic mug = 38

Administrator's note: No comment on the fact that temperature went up instead of down.

Although students who were taking algebra (our only measure of general school performance) tended to perform better than others, this was not necessarily the case. For example, the following pair of students were academically advanced but showed little understanding of the requirements of an experiment.

Poured some water into each beaker, put thermometers in, watched thermometers rise.

Poured water out of insulated cup. When I asked "Why?," they said, "Because it was too low (temperature) already."

Kept watching the remaining two cups and thermometers.

Poured more water into both cups, read both thermometers and said it was definitely the styrofoam cup.

The administrator commented: I don't know how they threw the insulated cup out so fast. They took temperatures of different cups at random. They guessed as to degree readings, rounded off to nearest temperature. The number of readings per cup varied. They did no recording of data. Amount of water was different in each cup. Although enthusiastic and engaged throughout, there was little apparent concern for accuracy.

Factors Associated with Success

Those who did well on the task:

- showed confidence
- remained engaged throughout until completion
- were careful about checking results
- were enthusiastic
- had done something like this before

Approximately 29 percent of the students said that they had participated in an experiment like this before. This experience did effect the performance of many. For example, the following pairs of students accomplished the task with efficiency. In both cases, the administrators noted that the students had previous experience with this kind of experiment. They noted that the students were confident and enthusiastic and were careful in checking results.

Example 1:

They filled beakers to 150 ml, placed thermometers in cups, make chart and then simultaneously poured the water into each container. Took first reading at 43 degrees once stabilized. Took another reading within seconds but did not refer to the stop watch. Continued to take readings. I discovered that Heather was taking 30 second reading. (Position of stop watch prevented my seeing this part of their work.) Continued to take readings at 30 second intervals for 5 minutes. Insulated cup held heat longest. chart made.

Example 2:

Thermometer in each cup—paper set up as data table.

150 ml of water in each beaker were transferred to cups, stopwatch for 30 seconds as initial reading. Read again at 1 minute, 2 minutes.

Partner admonishes selected student to keep thermometer in water. Added 3-minute and 4-minute readings.

Decided they should repeat to verify results (as per their teacher).

Insulated cup best, **styrofoam** next, **mug** worst

Second trial- **insulated** and **styrofoam** tied. **Mug** still worst.

On the other hand, many of those who were not successful appeared to enjoy the task. One administrator quotes a student as saying, "I liked this. It was fun to be chosen. I usually get left out." The administrator goes on to comment: "They had never been given a problem to figure out without being told what to do before. They have done experiments in science but not like this. They continued to discuss which cup they would buy as I write this."

Another wrote:

> These boys *loved* what they were doing!! At the end, they said they wished they could always take tests this way. When I asked them why, the selected student said, "Because it's more interesting and you actually learn stuff while you are being tested. It's more fun when it's not so important that you get one special answer."
>
> *The administrator goes on to comment:* "...strange, since selected student didn't conclude anything but what his initial guess was..even though he did waver slightly in the face of the data."

On the other hand, not all were so enthusiastic. Administrators rated approximately 30 percent of the students as neutral about the activity. Approximately the same percentage were seen to be uncertain about how to go about the problem. Although we did not look at the correlation between the two, anecdotal evidence suggests that the open-ended nature of the task exacerbated any anxiety that students may have felt when faced with an unfamiliar type of problems. In general, fewer students showed enthusiasm for **Insulation** than they did toward

Cubes and **Circuits**, both of which were more structured. An administrator's note about one pair of students further confirms this conjecture.

> "Students seemed afraid to use equipment, had no idea how to set up the experiment and wanted to be done as quickly as possible. Students really did not want to touch any equipment."

Summary of Results

Few students had a clear idea of what the task involved. Many took the instruction to "find out the cup that holds the heat the best" at face value and started to pour and measure without much consideration of the factors that might affect their results. Perhaps they did not even recognize the task as a "scientific experiment" for which certain procedures applied. In any case, few took the time to organize a plan or to identify the dependent and independent variables. It is doubtful that they even recognized their relevance.

Many students appeared to be concerned only with producing an answer. They did not recognize that *how* they went about the task was relevant or affected the validity of their answer. Administrators noted that only 35 percent of the students appeared to be concerned with accurate procedures, carefully checking their results. Although about half the students made some record of the results, there was very little attempt to use the data. Few made any attempts to examining it for patterns or anomalies, such as the different rates of heat loss of the various cups. In other words, the *process* of experimentation was of little concern to most students. Perhaps they did not see how a scientific approach could be applied to anything as prosaic as evaluating the efficiency of hot drink cups.

Students exhibited a variety of problems on this task. Briefly, they were:

- little clarification of the problem, planning of the experiment, or evaluation of the results.
- some inability to recognize the requirements of the problem, i.e., that it called for experimental controls and standardized procedures.
- little apparent understanding of the different types of variables involved.
- a lack of recognition that the purpose of the task was to record and interpret a relationship, not just produce an answer.
- a lack of understanding of the role of data in any investigation.
- some misconceptions about the transfer of heat and the conservation of volume.

Extension for Classroom Instruction

This is a change from the usual laboratory situation in which procedures are followed in order to produce the "right" result. Insulation asks students to investigate a real situation that is of some potential interest to them. Since we have no data on how these students would have performed in a more "text-book" type of experiment, we cannot compare. However, it does seem evident that, even if they are capable of using a scientific approach in their science classes, they do not recognize its applicability in other contexts. Although these students may have performed more "scientifically" in a situation that was clearly labelled as such, basic approaches to investigation should not be limited to the laboratory.

Educators contend that students perform poorly in science because science instruction is overloaded with terminology and factual knowledge to the exclusion of the practice of science. The usual response to such charges is a demand for more laboratory work. However, following text-book procedures to establish known outcomes (the most common form of lab work) is not really practicing science either. When students have been observed and questioned on what they are doing and why, they often show little understanding of the factors involved. Their main concern is getting the "right" answer (which, in many cases, they already know). Although they may be learning laboratory techniques, they are not necessarily learning to practice science. To practice science well requires a willingness to analyze problems, as well as the ability to follow certain procedures.

In addition, a scientific approach can be the subject for instruction—not in the abstract, as a bit of factual knowledge, but through activities. For example, the steps in text-book lab experiments can be cut out, scrambled, and given to students to order. Can they reconstruct the original order? What is the rationale? Could the experiment be done another way? This may lead to discussion and more thoughtful consideration of certain procedure. The knowledge gained may generalize beyond the class when students better understand the reasons for why their experiments are structured the way they are.

Other booklets in this series offer specific suggestions for investigations that can be used as the starting points for experiments. For example, **Popcorn Estimation** provides a series of group investigations involving weighing and measuring. **Circuits** list some tasks involving electricity. Beyond this, there are commercially produced books that give investigative, in contrast to confirmatory, experiments related to the concepts involved in **Insulation**.[1]

Such investigations, particularly when they are carried out by different small groups, can provide vehicles for practice in scientific thinking. Students can create their own questions, observe and record interesting phenomenon, make hypotheses and test them. The advantage of group work, aside from the obvious requirement that students must communicate with each other within the group, is that groups may vary in their results and interpretations. If such differences occur, they provide the opportunity for groups to defend their methods and justify their conclusions before discriminating peers. In such cases, not only do students begin to act like scientists, but they learn the necessity of recording accurately and communicating clearly. They also learn that their work has to be defensible, that a "right" answer is not necessarily right.

Adaptation for Evaluation

In discussing the aims of schools science, Rosalyn Yalow, a Nobel Laureate for her work in physics, stated:

> Science is not simply a collection of facts; it is a discipline of thinking about rational solutions to problems after establishing the basic facts derived from observations. It is hypothesizing from what is known to what might be, and attempting to test the hypotheses. This is the scientific process, and it can be broadly applied.[2]

In evaluating students' ability to apply the scientific process to problems, the correct answer is far less important than how students go about the task. Students' approach to the task also gives insight into their understanding of the concepts involved. For example, in **Insulation** it was evident from their actions that most students had little appreciation of the role of controlling variables in the experiment.

Those who have attempted to classify the kinds of behavior that should be assessed in science have suggested four distinct categories.

Planning:
- raising and clarifying problems
- designing investigations

Performing
- observing
- manipulating
- data gathering

Interpreting
- data handling
- making inferences and evaluations
- predicting and explaining

Communicating
- reporting
- justifying results

The evaluation checklist on the next page generally follows this format. It is similar to that used in the state-wide administration of the test. It primarily concerns the experimental process, how well students adapt what they know of the scientific approach to an actual problem. However, it also addresses how well students understand the problem itself and the concepts involved. In order to explore this, we asked the performance test administrators to discuss the task with the students after they had reached

---

[1] See Schools Council. **Change: Stage 3**. Milwaukee, WI: Macdonald-Raintree. 1973.

[2] "Industry and Science in the Early Grades." The Wingspread Journal, vol. 10, 2, 1988.

Checklist for Student Evaluation

Planning:

- asks for clarification when appropriate ___
- willing to evaluate different approaches ___
- identified dependent/independent variables ___

Execution:

- methodical approach ___
- attempts to verify results ___

- variable 1: water
  - controls for quantity ___
  - measures level of water ___

- variable 2: temperature
  - baseline established for each cup ___
  - number of thermometer readings ___

- variable 3: time
  - establishes initial reading for each cup ___
  - duration between initial and final readings ___

- gathering of data
  - records all readings ___

- use of equipment
  - waits for thermometer to stabilize ___
  - reads thermometer correctly ___
  - measures liquid precisely ___

Interpretation:

- recognizes trends ___
- notes anomalies ___
- accounts for outcomes ___

Communication:

- gives clear description of outcome ___
- reports systematically ___
- displays appropriately ___

Attitude:

- willing to cooperate ___
- listens to others' ideas and responds ___
- confident about task ___
- confident about use of equipment ___
- remains engaged ___

their conclusions. We also asked administrators to keep any written records that students had made. These various sources of information complement each other and increased our understanding of student achievement.

Although the evaluation of students' performance ability requires actives experimentation, other aspects of the ability to engage in science can be measured in other ways.

For example:

raw data from an experiment can be presented and students can be asked to reduce, find patterns and make inferences (Interpreting):

the steps of an experiment can be outlined and students can be asked to describe the initial problem, including the important variables (Planning);

an experiment can be described and students can be asked for a critique of the methods used and/or the conclusions derived (Planning/Interpretation).

The most important requirement is that evaluation tasks allow latitude for different methods, ideas and solutions and that they challenge students in ways that have some meaning to their lives.

---

designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Background Summary," "Circuits," "Cubes," "Math Town," "Playground," and "Popcorn Estimation."

BEYOND Paper & Pencil

Popcorn Estimation

Introduction

To solve most math problems one has to be able to use proportional reasoning. It is basic to all operations that relate mathematics to the physical world, such as measuring, counting, estimating and calculation. It is a necessary requirement in algebra, geometry, calculus, statistics, chemistry, biology and physics. In fact, Piaget in his investigations of cognitive development regarded proportional reasoning as a prime indicator of intellectual growth.

Popcorn Estimation, which asked students to estimate the amount of popcorn kernels in a container, was designed to evaluate students' ability to apply proportional reasoning to a measurement problem. It required students to establish a ratio between quantity and some unit of measurement and then to use that ratio to determine an unknown. In the process, we observed students' strategies in going about a task, their use of equipment, their understanding of measurement units, and their ability to calculate.

Description

Popcorn Estimation was the least structured and the most open-ended of the problems presented in this series. Students were told to use whatever they wanted from a set of materials to estimate the number of kernels in a container.

> "The only thing that you can't do is count all the kernels in the container. The closest estimate to the real number would be the best estimate for me."

The role of the administrator was limited to that of an observer and recorder. When the students arrived at their solution, the administrators checked the accuracy of their own observations by asking the students to describe their method. Students were then asked to suggest an alternative strategy for solution. If appropriate (i.e., if students were able to describe a second method and if there was time), they were asked to repeat the task using their proposed method. Twenty-eight percent of fourth graders and 66 percent of eighth graders did this.

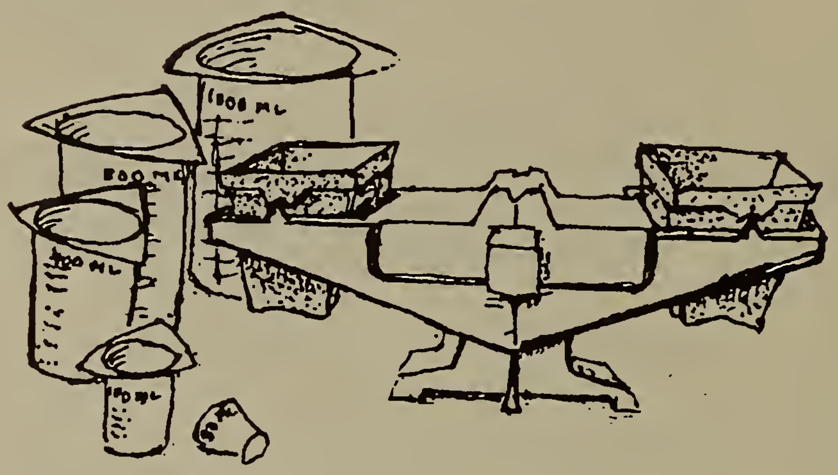
Massachusetts
Educational
Assessment
Program

Massachusetts Department of Education **January 1990**

In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

Materials:

- kernels of popcorn in original container
- scoop
- 30 ml medicine cup
- set of containers of different sizes marked in millilitres
- balance with pans
- a set of weights totalling 120 grams



Formulating a Plan

Before beginning the task, students were asked to discuss a plan of action which would form the basis for their work. It was expected that questions would be asked and different approaches considered; however, in most cases, students plunged right in. As a consequence, some misunderstood the requirements of the task. A few interpreted the administrator's instruction to mean that NO kernels should be counted. Others believe that the scale MUST be used. Although these misconceptions were usually corrected in the course of the sessions, there was little active seeking for clarification before students began the task.

This lack of reflection may have been caused by the testing situation itself; students may have seen the need for a prompt response. It may also have been caused by their unfamiliarity with this type of practical activity. When asked, only 10-15 percent of the students replied that they had ever done anything like this before. It was obvious that, in many cases such as the one described below, the problem itself functioned as a learning experience for the students involved. In any event, 55 percent of fourth graders and 43 percent of eighth graders changed their strategy as they progressed.

First Attempt: Started filling one side of scale with kernels (uncertain how much to fill balance container). Put all weights on other side. Kernels were too heavy, so took some out. No idea how this process would help solve problem.

Second Attempt: Poured all kernels in largest measuring container. Read 900 ml on container. Realized that 900 ml wasn't the answer to the question.

Third Attempt: Counted kernels needed to fill one scoop (100). Took continuous scoops until all kernels were used. No accurate measuring of amount in each scoop. Counted by hundreds as they added each scoop to get answer: 3200.

Selecting a Strategy

In addition to guessing, which was the strategy chosen by approximately 6-8 percent of the students at both grades, there were three basic strategies that could be used to solve the problem. All involved proportional reasoning, although they differed in their level of abstraction.

Strategy A. Students could count the quantities in a container and then determine how many containers it would take to hold all the kernels.

This is the simplest procedure that will lead to reasonable results and it was followed by approximately 30 percent of the fourth graders and a half of the eighth graders. Its analogy in a word problem would be: if one cup holds 10 kernels, how many will 10 cups hold? Carrying out the procedure involves simple physical actions. After establishing the initial units (number of kernels in a container), all that must be done is to count the number of containers that is equal to the total quantity of kernels.

The procedure could be relatively inefficient and/or inaccurate, depending upon the size of the container chosen. Although one could get a fairly accurate count of kernels in a small container, such as the medicine cup (which had a capacity of 30 ml liquid and held approximately 280 kernels), from 20-25 cups were needed to measure out the total number of kernels. Not surprisingly, many students lost track or became careless by the end. On the other hand, the larger containers held so many kernels that an accurate counting was almost impossible and most certainly a long and tedious process. Although far fewer scoops were needed to measure the total amount of kernels, the unit of measurement was bound to be inaccurate.

Strategy B. Students could establish a proportional relationship between the quantity of kernels and their volume. Since the total volume could easily be determined by measuring, they had to solve for the total quantity of kernels only.

The containers were marked in milliliters, and two of the larger ones had sufficient capacity for all the kernels. Consequently, students could easily find the volume measure for the entire lot of kernels. Having determined this, they could establish the ratio between a smaller volume (say, 10 ml) and the number of kernels. Once students established this relationship, their main task was to apply proportional reasoning to find the number of kernels that would be equivalent to the total volume. The model for this strategy is:

$$\frac{\mathrm{ml}}{\mathrm{k}} = \frac{10}{100} : \frac{700}{?} \quad \text{(known)}$$

Strategy C. Students could establish a proportional relationship between quantity of kernels and their mass. Since the mass could not be determined by the number of weights available, they had to solve or both mass and total quantity of kernels.

Although superficially the same as the problem above, this strategy involving weighing was far more difficult because there are two unknowns. Not only did the total number of kernels have to be found, but the number of grams representing the total mass had to be established. Whereas, in the earlier strategy, the total volume could easily be found by pouring the kernels into one of the large calibrated containers, there were not enough weights to balance all the kernels. In order to carry out this plan successfully, students had to have a clear enough plan to carry them through the additional procedure of repeated weighing of the kernels.

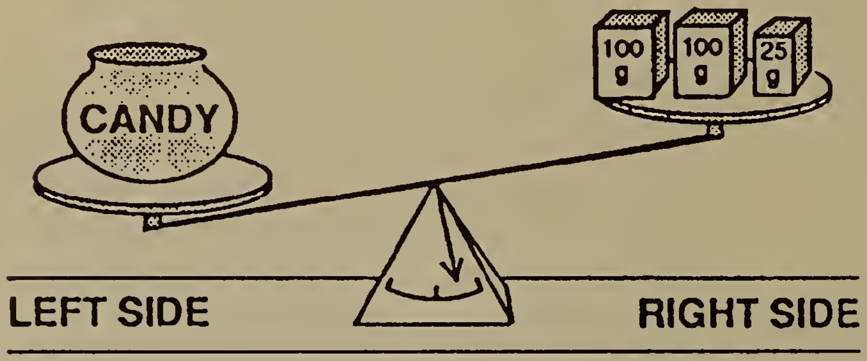
$$\frac{\mathrm{g}}{\mathrm{k}} = \frac{2}{24} : \frac{?}{?}$$

Of course, the implications of these strategies were not apparent to students as they began the task. Nor, for the administrators, was there a preconceived hierarchy of strategies by which to judge students. The more sophisticated strategies entailed hazards that are beyond the bounds of most fourth and many eighth grade students to cope with. The purpose of the administrators' observations was to see how students perceived the requirements of the task and how they carried out their plans.

Fourth Grade Results

Possibly intrigued by the opportunity to use the balance, approximately half the fourth grade students suggested a procedure involving weighing. Most were vague about their plan: Put popcorn in one pan, weights in the other. However, 20 percent of the children who decided to use the balance voiced a specific plan such as, "Weigh a small amount to find out how many kernels are in one gram—then add them up."

Unfortunately, the conceptual and practical demands of the weighing strategy thwarted the plans of almost all. Even those who went so far as to establish some sort of equivalency (e.g., 10 kernels = 1 gram; 60 kernels = 6 grams) were eventually defeated by the necessity of determining the mass of **all** the popcorn kernels and then relating this back to their initial equivalency. Many became confused by the additional complication of the calibrations on the measuring containers—confusing milliliters for grams. Others became mesmerized by the balances themselves, forgetting the purpose of the task in their concentration on achieving a perfect balance. In fact, the balance itself proved to be a problem for many children. Thirty-one percent of fourth graders (and 39 percent of eighth graders) were observed to add weights to the heavier side or remove them from the lighter. This was in sharp contrast to their generally correct responses to a multiple-choice question asking for their knowledge about balancing. They obviously found it hard to apply what they appeared to know.



To which side of the balance must more weight be ADDED to make the two sides balanced?

		Gr. 4	Gr. 8
A.	right side	70	86
B.	left side	21	10
C.	neither side	3	2
D.	cannot tell	5	2

• • •

The children described below are typical. Intrigued by the act of balancing, they did not seem to understand either the process of weighing or of measurement. Nor did they seem to relate their activity to the purpose of the task—determining the number of kernels. Instead, they assumed that the calibrations on the measuring containers supplied the answer. Note also, when they added fresh quantities to each pan, they did not seem to recognize that they were adjusting their previous total. Instead, they treated each new total as a separate entity and totaled all numbers that had been noted previously.

poured kernels into both containers of balance

tried to make even, did not use all the kernels

added kernels to both sides until even (added to wrong side several times)

placed kernels from one side of balance into large beaker (recorded volume 300 ml) and replaced it in the balance pan.

placed kernels from other side into another beaker (recorded volume 300 ml) and replaced in balance pan.

recorded total: 600 ml

continued adding kernels to both sides, trying to keep balanced. (Became very involved in making scale balance perfectly, adding one kernel at a time to each side)

when balanced, placed kernels in each pan and recorded volume (350 ml) for each

added 350 + 350 to get 700

replaced kernels and continued adding to both sides

repeated volume measurements

added all volumes to get answer: 2600

(300 + 300 + 600 + 350 + 350 + 700 = 2600)

Administrator's comment: These two had little grasp of the problem. They spent most of the time reinforcing their learning about how balances work.

Remembering which units they were working with was a common problem with these younger students, particularly when they chose a weighing strategy. The conceptual demands required by this strategy appeared to distract students, even those who appear to have an otherwise clear plan for solution, such as the students described below:

---

weighed 250 ml container—13 g

filled it with corn and tried to balance with gram weights—did not have enough

moved to next smaller container—weighed 11 g

filled with corn—weighed 105 g

subtracted weight of empty cup from total—94 g

looked at remaining kernels and guessed that there are about 4 more small containers left—raw estimation

they then took 94 X 5 and got 452

they then said they need to add an extra zero because 452 did not sound right

4520.

The fact that these students carefully weighed and subtracted the weight of the containers indicates a concern for accuracy. That they were not put off their plan by the fact that there were not enough weights to balance also speaks well for their organization. Finally, their general plan was good: if 1 : 94, then 5 : 452. However, they appeared to be so interested in the process of weighing that they equated the number of grams with the number of kernels. Their final response suggests that they recognized that their answer (452) was incorrect, but did not realize why.

• • •

The students described below also used the balance, but not for weighing. Instead, they used it to split the quantity of kernels before adopting a counting strategy.

---

put in approximate cups of kernels into each side of scale

increased cups on each side of scale and balanced this out

now figured out the number of cups using paper to keep track of what they were doing

kept on measuring kernels and balanced scales. came up with 8 1/4 cups

(I reminded them to give the number of kernels.)

used small container and counted out by tens. figured out the number of kernels to fill 30 ml container—about 200 kernels

figured that 3 containers would equal 1 cup—600 kernels

decided there were about 8 cups in each side of scale

600 kernels times 8 = 4800. 4800 + 4800 = 9600.

Administrator's comment: Students immediately knew how to attack this task. They were able to use the scale correctly in their problem solving. They were not always accurate in their measuring of cups, balancing of sides or even counting of the kernels to fill one 30 ml container. They did, however, attempt to use equivalent measures and arrived at a logical conclusion. They were also able to think quickly of another way to solve this problem. Calculations done were correct.

A few fourth grade students did succeed with a weighing strategy as illustrated below:

---

measured kernels to 250 ml mark on container

placed kernels in one pan of balance with weights in other. did not work

decided to count 50 kernels. placed these in balance pan and balanced with 5 g

placed 120 g in—balanced with kernels

divided 120 by 5 = 24

multiplied 24 X 50 = 1200 kernels

$$\frac{k}{g} = \frac{50}{5} = \frac{k}{120} \quad (1200)$$

continued to balance 1200 kernels, repeating procedure 4 times (4800)

remainder was balanced by 95 g. divided by 5 = 19

19 X 50 kernels = 950

4800 + 950 = 5750 kernels.

$$\frac{50}{5} = \frac{k}{95} \quad (950)$$

These students showed a mathematical sophistication that is unusual, not only among the set of students tested here but among those reported elsewhere. Using the ratio of 5 g to 50 kernels, they were able to calculate the number of kernels that would weigh 120 grams. They then used this relationship to weigh and count the entire quantity of kernels.

•••

Despite the fact that the volume strategy (Strategy B) presented less difficulties than weighing, few children who used it were able to work through to a solution. Partially, this may have reflected the difficulties due to proportional reasoning. However, it was clear that many had little experience with the measurement of volume or understanding of what the ml calibrations on the containers stood for. For example, many who poured the kernels into one of the containers interpreted the milliliter calibrations as the number of kernels. Others seemed clearly adrift, making calculations with whatever numbers appeared, without much concern for the units or the type of question asked. The following explanation describes such confusion:

> "We found out how much all the kernels weighed. Then we found out their height in the container. Then we divided the height [milliliters] into weight to get 100 remainder 342. Then we added 100 to the remainder to get 442."

Even children who recognized the proportional nature of the task experienced difficulty in bringing the problem to a successful conclusion. When computational errors were made, they seemed to have little appreciation for the need to check their calculations against a reality benchmark. For example, one pair of students who estimated that there were approximately 25 kernels to one milliliter, multiplied by 700 (700 milliliters, the volume of all the kernels) and obtained the figure: 1,750,000. The students neither checked the actual count (which was closer to 8) nor did they appear to be fazed by the size of their result. In addition to the conceptual difficulty of visualizing large numbers (without an idea of what 175,000 kernels look like, it is difficult to imagine 1,750,000), there was the added handicap of lack of ease with calculations themselves. In general, children avoided division and preferred addition to multiplication in working out their answers.

Eighth Grade Examples

Eighth graders showed greater maturity and ability to keep to their plan. Fewer chose weighing initially. Most used the volume strategy (Strategy B) or counting (Strategy A).

The two pairs of students described below both used the balance, but in different ways. The first pair used a more formal approach to calculate the number of kernels; while the second used the balance to measure the volume of kernels but basically adopted a counting approach.

Pair 1:

tried weighing 1 kernel—no reaction

tried weighing 25 kernels—2 grams

they then poured nearly all in but didn't have enough weights to mass it. Decided to half it; reduced approximately 800 ml to 400; halved it again to 200, then to 100.

100 kernels weighed 73 grams. If this is 1/8 of total, all kernels = 584 grams.

2 grams divided by 25 = .08 g per kernel

now divided 584 by .08 = 7300.

$$\frac{k}{g} = \frac{1}{.08} = \frac{x}{584} \quad (7300)$$

Pair 2:

found out that 13 kernels equals 1 gram—counted out 100 kernels and added to the 13 in pan, found out that 113 kernels weighs 11 grams—added another 87 kernels to make 200 which weighs 22 g.

emptied 200 into another container

weighed out another 22 g worth and emptied into storage

added another 22 gr masses to make 44 g of masses and balanced—400 kernels

400 kernels (partner keeps tally on paper)

added 22 g masses to make 66 gr of masses—balanced with 600 kernels

repeated this step three more times (4 lots of 600)

procedure change: decided they had done approximately half of kernels—put these in one pan, takes weights out, pours remainder in other pan—not quite equal

makes them balance correctly by removing some kernels.

weighs leftover exactly.

Since 1 g = 13 kernels and leftover weigh 18 g, 13 X 18 = 234

2600 + 2600 = 5800 + 234 = 6034 (calculation error)

The following pairs of students were more successful and were typical of those who used a proportional approach to volume (Strategy 2). One pair integrated this approach with weighing; the other counted initially.

Pair 1:

Weighted 8 kernels to be 1 gram.

Filled 1 side of scale with kernels. Started to weigh.

Decided that there were not enough weights to weigh that many, so started over.

Poured 100 ml of kernels into one side of scale and weighed them—94 grams.

Multiplied 8 (kernels) X 94 (grams) and got 752 kernels in 100 ml.

Measured all the popcorn as 900 ml.

Multiplied 9 X 752 and got 6768.

Added 8 (original amount not measured) to 6768 and got 6776.

$$\frac{ml}{k} = \frac{100}{754} = \frac{900}{k} \quad (6768)$$

Pair 2:

Put all the kernels in one container—filled to 700 ml.

They filled the 100 ml cup to 70 ml. This would be 1/10 of the kernels.

Together they carefully counted the kernels in the 70 ml. There were 513 kernels in 70 ml, so they multiplied by 10 and got 5130.

One boy suggests that they fill a 100 ml cup and multiply by 7. The other nixed it.

$$\frac{ml}{k} = \frac{70}{513} = \frac{700}{k} \quad (5130)$$

Although the milliliter calibrations on the measuring cups were probably unfamiliar to most fourth graders, many eighth graders also found them confusing. However, this may have only exacerbated a more fundamental problem that students had with the concept of measurement—that measurement is basically a system of relationships among different units. Many did not recognize a need for establishing a unit of measure for one kernel, but relied on a relationship between mass and volume. In doing so, they confused the units and lost track of the purpose of the task. For example, the students described below showed little understanding of density, although they knew that there was a formula involved.

> Used large container to measure 750 ml—packed down
>
> Attempted to weigh 750 ml—wondered if it would fit into scale container—packed kernels down to fit
>
> Placed weights into opposite side—counting as they went—partner was sure they did not have enough weights but selected student kept on counting
>
> Realized that they did not have enough weights
>
> Decided to weigh 100 ml—poured out 250 ml from scale—followed same weight procedure
>
> 100 ml weighed 103 grams
>
> Found density of 1 kernel by dividing mass (103 grams) by volume (100 ml).
>
> Kernel's density was 1.03.
>
> Multiplied 1.03 X 700 ml = 721 kernels

Since this answer was very different from the one arrived at during their first attempt at the task, the administrator asked why they arrived at different estimates.

> "You can't come out exactly the same each time. Each time you weigh things they are not exactly the same. Weighing salt and pepper is different because of the different densities."

Confusion about the concept of measurement was pervasive. In many students, it took the form of mindless calculations, as in the descriptions below:

> ***Pair 1:***
>
> Counted 42 kernels which equals 5 ml.
>
> Multiplied 42 X 5 = 210.
>
> Stopped. Change procedure.
>
> New procedure is to divide whatever total ml of kernels is by 210 (grams mentioned)
>
> Total millilitres of kernels found to be 700 ml.
>
> Divided 700 by 210 and got 3.33.
>
> "Can't be—we're a little off!"
>
> Change procedure: multiply 700 by 210 = 147,000 (approximately equal to)

> ***Pair 2:***
>
> Began to counter kernels as he filled the 100 ml container.
>
> Decided would be there all day. Poured into 50 ml container—416 kernels.
>
> Poured rest of kernels into large container. They filled to 700 ml mark.
>
> Divided 700 by 416. "No, 700 by 50." (Selected student saw error of first suggestion.) Answer: 14.
>
> Then divided 700 by 416. Answer: 1.6.
>
> Went back to answer of 14 and multiplied by 416. Got 5824.
>
> Thought this answer was too low. Some discussion. Selected student was going to divide 700 by 416 and multiply answer by 50. Gave up on the division step.

Whereas the students described above recognized that some of their solutions made no sense, once others had begun to calculate, they appeared to have no benchmark for the reasonableness of their answers. This was particularly apparent when students arrived at two radically different answers for their first and second attempts. The description below illustrates this, as well as giving us a glimpse at these students' understanding of the meaning of estimation.

Students initial answer was 14,000. They arrived at this by filling the 20 ml container with 20 kernels. They then determined that there was a total of 700 ml. "Twenty seeds filled 20 ml and there is a total of 700 ml, therefore 14,000 seeds."

In their second attempt, the students counted 162 kernels that formed the bottom layer of a container. These rose to the 20 ml mark on the container. Thereupon they multiplied 162 by 700 and arrived at the figure of 113,000.

When asked why their numbers were so different, they replied that they had underestimated the first time. "The second answer was best because the first time we didn't count the seeds but estimated. The second time we counted."

• • •

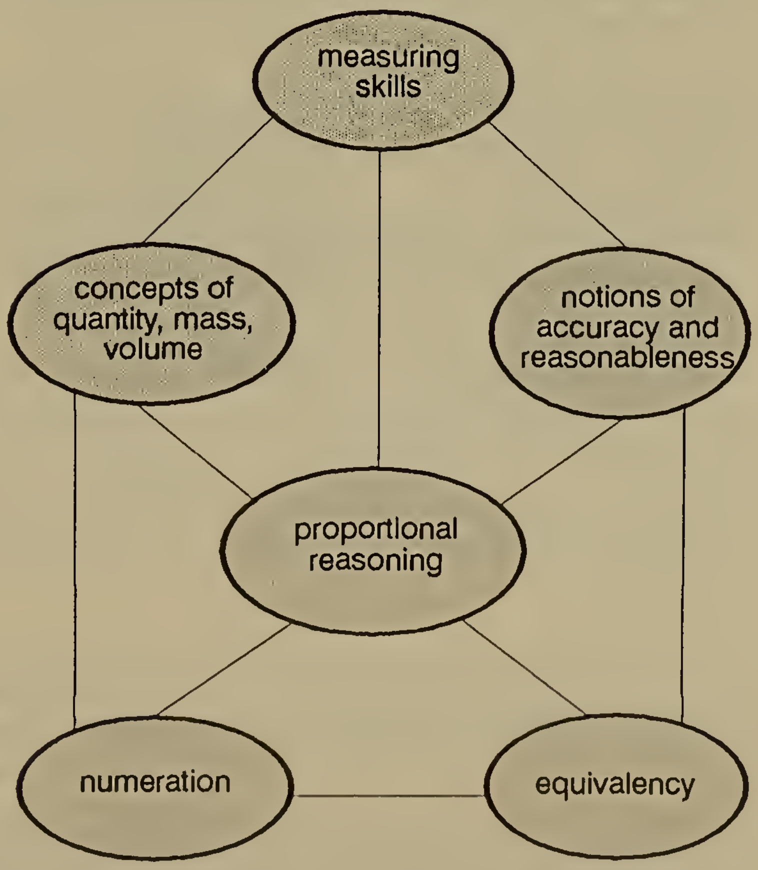
Summary

Perhaps it is a general rule that the more open-ended the task, the more revealing it is in terms of students' attitudes, strategies, and understandings. Certainly, this simple task proved extremely challenging for most fourth graders and many eighth graders. The lack of structure or explicit direction, the attraction of the balance, the constraints that were not initially obvious—all contributed to the difficulties that students experienced. On the other hand, such problems are important for both evaluation and instruction. Through them, we learn how well students can clarify problems, make plans and use equipment. In addition, we can observe how well they understand the concepts involved.

This performance task shows a variety of problems that children experienced. Briefly, they were:

- a lack of familiarity with using a balance
- a lack of understanding of the concept of measurement as a relationship between units
- unfamiliarity with metric units (e.g., millilitres)
- a tendency to retreat to meaningless numerical computations that were divorced from any referrents
- little evaluation of the results in terms of reasonableness
- misconceptions about the requirements of a good estimate

Of course, many of these difficulties stem from a lack of maturity and, certainly, there were much more prevalent at the fourth, in contrast, to the eighth grade. However, they are not unaffected by lack of experience, and most students appeared to have little familiarity with this kind of problem.


measuring skills
concepts of quantity, mass, volume
notions of accuracy and reasonableness
proportional reasoning
numeration
equivalency

Extensions into Classroom Instruction

There is no doubt that children have some intuitive understanding of proportional reasoning. Whenever they solve problems by thinking "for every ...," they are displaying this. With maturity, they begin to recognize and coordinate more than one variable in their mathematical thinking. However, the common tendency of textbooks to concentrate on numerical calculations at the expense of concepts leaves students with an incomplete understanding of what ratios mean and how they are used. Studies suggest that "without the units of measurement, ratios loose their meaning and students have no way to assess the reasonableness of their manipulations and results."[1]

Measurement, which is basically the establishing of relationships, is an excellent context for developing the kind of understanding needed for proportional thinking. It constitutes a practical context in which students can begin to explore equivalency of units and the need for invariance. These experiences with physical operations lay the groundwork for later abstractions and more formal calculations.

The instructional extension of the ideas contained in **Popcorn Estimation** concentrate on equivalency and measurement, rather than on number alone, because these concepts are fundamental to the development of proportional reasoning and general problem solving.

Volume. Give the students a collection of containers, some beakers, a balance with weights and some sand. Which containers hold the most sand? Which holds the least? Do any hold the same amount? Do they hold the same amount of water as they do sand?

Equivalency. Collect some blocks, stones, or rods of the same and different size. (Tangrams will also do.) See how many different combinations will balance. Record your results (drawing is best). See what patterns you can find. Use the balance to try to order the objects from heaviest to lightest.

Volume/Displacement. Collect a group of different-shaped bottles. Measure the volume of one, then estimate how much the others will hold. Were they good estimates? Make a record of how much each holds and order the bottles in terms of their volume. Now see how much each bottle weighs. Record the results and compare. Is this what you would expect? Can you find out how much plastic or glass is used to make different bottles?

Measurement Accuracy. Have different groups of students measure the same set of things and compare results. How near are the answers? How near would you expect them to be? Are any of them very different? What do you think is the right answer?

Measuring Skills. Have the students experiment with inventing and making weighing machines for themselves. These may involve elastic or spring. Other balances will work on a lever or pivot principle. What are the best ways to weigh different things? How would you weigh the milk in a bottle? How would you work out an estimate for the weight of a seed? A pin? A flower petal? Can you work out a technique for estimating the weight of the principal's desk? the school?

Density. Collect a set of blocks or different kinds of bricks that are as similar in size as possible. Compare the mass of each and order them by weight. Now work out a figure for the density of each by finding the mass per unit of volume. Use different materials (styrofoam, clay, wood) to make some shapes that have the same mass but different densities. Make some that are of the same size but have different densities.

Estimation. What is the difference between guessing, estimation and rounding off? When would you use each? When **wouldn't** you use each? What makes a good estimation? How would you go about estimating the length of the school yard? The number of left-handed people in your school? How much money the people in your class spend on candy each week?

---

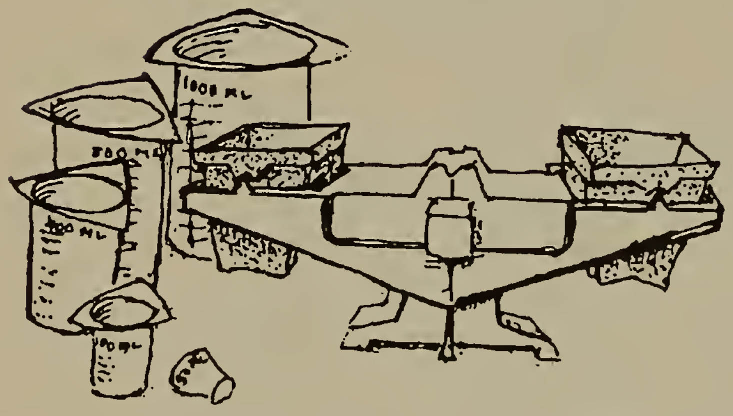
[1] The discussion of proportional reasoning was informed by Romberg, T., S. Lamon and E.A. Zarinnia, "The essential features of the mathematical domain of ratio and proportion," National Center for Research in Mathematical Sciences Education, University of Wisconsin, Madison, October 1988.

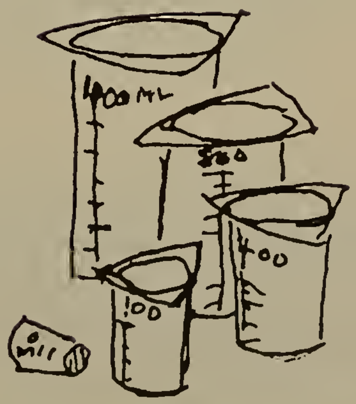
Equivalent Units. Give the students a very large, a very small glass jar and some marbles. Use whatever you know about measurement to find out how many marbles the large jar will hold. You can use sand, water, measuring cups, rulers, etc. to try to find out.

Finally, students can be encouraged to engage in more general problem solving strategies, such as making a plan of action. Although "playing around," getting a feel for the problem, and experimenting with some tentative approaches are important to problem solving, eventually successful problem solvers decided on a plan of action. One way for students to practice thinking about such plans is to think in terms of a series of steps. Although the eventual aim is for students to direct themselves, it is easier to practice by giving directions to others. Start with some everyday practical activities that children are familiar with, such as making a peanut butter sandwich. Next move on to less familiar ones, such as describing a shape for another child to draw. Writing a set of directions not only calls for clear writing, it calls for clear, sequential thinking—in other words, a plan.

Adaptation for Evaluation

Problems such as **Popcorn Estimation** are useful for instruction. However, they also yield important information about what students can do and how they understand the concepts involved. General impressions or final products are not sufficient for this kind of information. Our own analysis of the responses of students showed that, although a wildly inaccurate answer was a sign that students had not understood the task, a correct one could not be interpreted as understanding. Focused observations gave us the data that we needed to come to any generalizations about what students know and understand. The chart on the next page attempts to set out a framework for such observations. It can be adapted for any of the activities given above. It need not be completed for every student on every occasion, but the habit of observing and recording such information permits the kind of focused instruction that is the mark of good teaching.





Observation	U/N	Comment/Example
Measuring Skills		
measures accurately		
uses correct unit for mass		
used correct unit for volume		
Proportional Reasoning		
level of abstraction:		
counting		
one unknown (volume)		
two unknowns (grams)		
mixture of above		
Equivalency		
adds/subtracts to balance appropriately		
relates equivalent measures (e.g., grams/mls)		
keeps ratios constant		
Numeration		
computes correctly		
multiplication		
division		
addition		
uses multiplication in place of addition		
recognizes correct referrant		
checks for reasonableness		
distinguishes between estimation and "rounding off"		
Planning/Execution		
asks for clarification when appropriate		
willing to evaluate different approaches		
orderly procedure		
makes written record		
checks results		
Attitude		
works with other student		
dominates other student(s)		
confident about task		
confident about use of equipment		
remains engaged		
enthusiastic		

---

designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Background Summary," "Circuits," "Cubes," "Insulation," "Math Town," and "Playground."

BEYOND Paper & Pencil

Cubes

Introduction

Approximately 70 percent of eighth grade teachers state that they place considerable emphasis on the computation of fractions, decimals, and percentages. Eighteen percent state that they emphasize mathematical work using plane and solid figures; while 7 percent stress spatial visualization in their instruction. This balance of effort may reflect the norm, but it should not be interpreted as the ideal. Visualizing and working with geometric figures should be a prominent part of the primary and middle school curriculum. They are the vehicles for mathematical concepts that students will meet more formally in courses such as calculus and geometry. Equally important, at every level they are useful in making number relationships more meaningful to children.

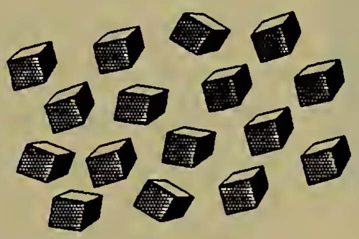
In **Cubes** we investigated students' ability to work with solid figures to solve problems involving both number relations and measurement. We also looked at how students' strategy was related to their success.

Problem 1

Students were asked to determine how many different rectangular solids could be made with 16 cubes. Administrators demonstrated a rectangular solid if necessary. They also gave the students a chart and suggested that they use it to keep track of the different shapes.

Materials: 16 wooden cubes

a chart with columns labelled **height, length, width.**



height	length	width

Massachusetts Educational Assessment Program

Massachusetts Department of Education **January 1990**

In the spring of 1989 over 2000 pairs of fourth and eighth grade students were assessed on their ability to apply mathematical and scientific concepts. This series of reports describes and discusses the results of these assessments. Prepared by Elizabeth Badger, Brenda Thomas, and Elizabeth McCormack.

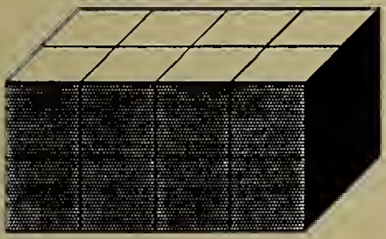
Focus

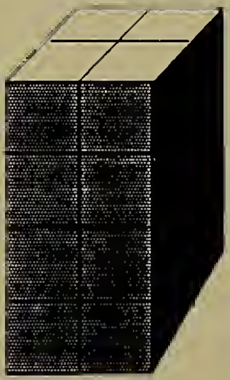
This task evaluated students' understanding of rectangular solids and their ability to recognize a physical representation of factor relationships. The chart, with its labelled columns, was given as a prompt to recognizing this relationship, as well as an aid in organizing the results. If students did not recognize the similarities between the solids that they were constructing and their numerical properties, were they able to abstract the generalities from the chart? We also observed how students viewed the various dimensions of a rectangular solid. Did they recognize that height, width, and length are not absolute dimensions but vary according to the orientation of the figure? For example, did they recognize that in a 2x2x4 cube the 4 can correctly refer to width or length?

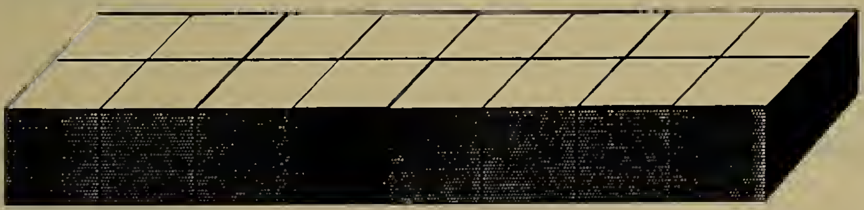
Results

Seventy percent of students produced five different rectangular solids and a few at both grade levels went on to make a complete set.

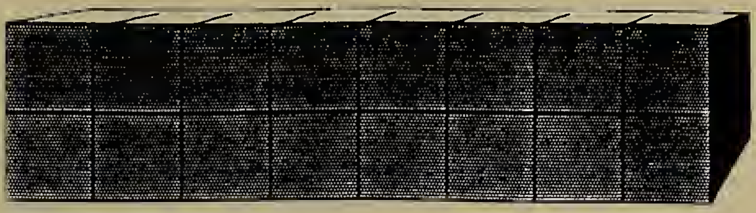
The most popular shapes, produced by almost all, were two cubes wide. These may have appeared to be prototypical rectangular solids to the students.

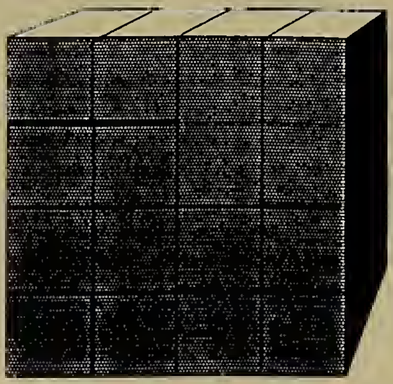

2x4x2


4x2x2

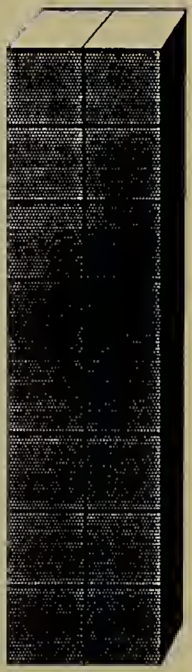

1x8x2

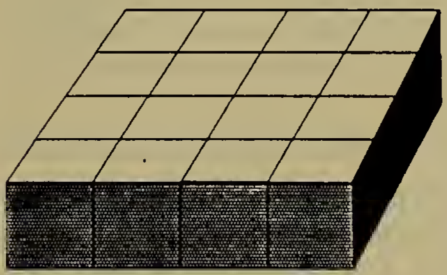
It was only after other shapes had been explored that students began to produce rectangular solids that were one cube wide, and two-thirds of the students did not recognize them at all.


2x8x1

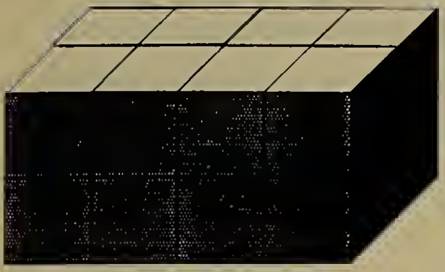

4x4x1

Approximately 35 percent of fourth graders and almost half of the eighth graders rejected a 1x4x4 solid on the grounds that it was a square rather than a rectangular solid. That it was rejected by more of the older students may reflect their misunderstanding of terminology. We have noted in other reports that students seldom recognize squares as instances of rectangles. Consequently, we surmise that the words, *rectangular* solid, may have prompted them to reject anything that could be classified as a square.


8x2x1

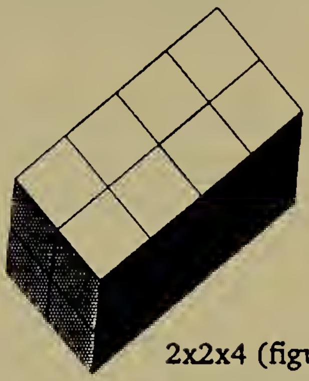

1x4x4

Other than as an aid in record-keeping, the charts provided little help in discovering the entire set of solids that could be made from the 16 cubes. Only 10 percent of fourth graders and a quarter of eighth graders appeared to realize that the dimensions of the solids had to be factors of 16, and almost half the fourth and 20 percent of the eighth graders continued to try to make rectangular solids using 3's or other non-factors of 16



2x4x2 (figure x)

throughout the problem. Few recognized the pattern that began to emerge on the chart as they recorded the dimensions, nor did students appear to understand the relationship that exists between length and width. For example, although 72 percent of the students made and recorded 2x4x2 as a solid (see fig. x), only 9 percent recorded 2x2x4 (fig. y); while 75 percent of students recorded 1x8x2, only 11 percent gave 1x2x8. Although such distinctions would be difficult to see visually, if the students had understood the basic relationship between the numbers, they would have automatically recorded both, once one had been constructed.



2x2x4 (figure y)

For example:

height	length	width
2	4	2
2	2	4
1	8	2
1	2	8

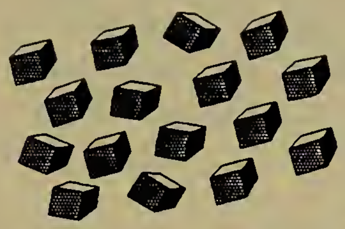
Some students may have learned throughout the task. When they had completed the problem eighth graders were asked if they recognized a pattern on the chart. Seventy-five percent answered that they did and almost 40 percent stated that they could now see how they might have come up with the shapes using only the chart. On the other hand, few recognized it without such specific prompting.

Problem 2

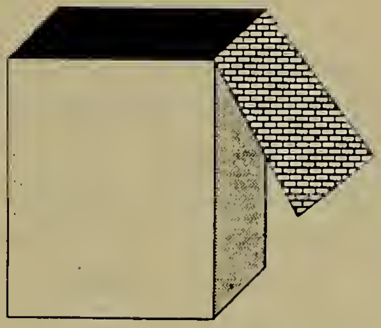
The students were given a container and asked to find out the maximum number of cubes that would fit.

Materials: 16 wooden cubes

container with a capacity of 96 cubes



16 cubes



Container

Focus

This task investigated students' understanding of the three-dimensional nature of volume. Not enough cubes were given for them to actually fill the container and count. They had to determine the dimensions and apply some sort of calculation. Not only did we look at the results obtained, but we noted the sophistication of students' strategies.

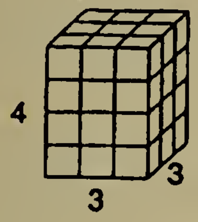
Results

Percentages correct in determining the number of cubes

	Grade 4 n = 138	Grade 8 n = 125
Boys	44	64
Girls	62	77

Girls were generally more successful at this task than boys. In fact, fourth grade girls performed at the same level as eighth grade boys. Most of the students who gave the correct number of cubes knew how to calculate volume. They had performed similar calculations on a previously administered multiple-choice question in which the dimensions were given (see below). However, an additional number of girls who did not answer the multiple-choice correctly seemed to benefit from being able to handle the actual materials. These girls filled or estimated a layer

of cubes and then multiplied the number of layers by the height (16 x 6) to arrive at the correct answer.



The large block above is made up of 4 layers of small cubes. Each layer has 3 rows with 3 cubes in each row. How many small cubes are there?

Percentages correct	Grade 4	Grade 8
Boys	46	78
Girls	38	76

Among fourth graders who were incorrect, 15 percent confused area and volume. Most of the other errors were the result of miscounting or incorrect estimation, perhaps due to a more basic lack of procedural knowledge. Fewer eighth graders (8 percent) confused area and volume; most of their errors were due to carelessness, particularly among boys. For example, 22 percent of eighth grade boys made some errors in counting the number of cubes. Again, these results can be contrasted with those achieved on the multiple-choice.

Problems 3, 4, and 5

The administrator presented Shape A, B, and C in turn (Shape C was not presented to Grade Four), and asked "How many more cubes are needed to fill in this rectangular solid without making it any bigger?" The original 16 cubes were also available as a concrete aid for students.

Focus

These three problems also involved students' understanding of volume; however, the problems were more abstract and were certainly less familiar to the students.

In these problems, with their skeletal structures, the students had to imagine what the completed rectangular solids would look like before they even began the task of calculating the missing cubes. Furthermore, the increasing complexity of the figures called for increasingly more elegant methods.

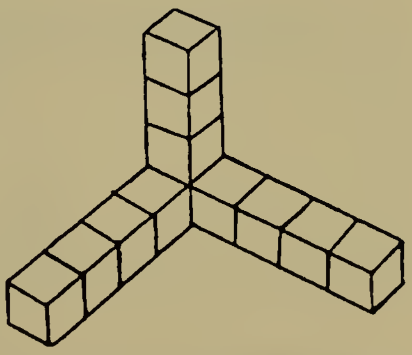
For example, the missing cubes in Shape A could be determined by placing a layer of cubes at the bottom and then estimating the number of layers needed to make a complete solid. In other words, the answer could be calculated by counting the number missing. However, the complexity of Shape B makes this strategy cumbersome. Cubes appear at different points, interfering with the easy placement of "counting cubes." More refined strategies, such as determining the outside dimensions and multiplying, are more successful than the counting strategy that suffices for Shape A. Finally, the half and quarter cubes that are missing in Shape C call for different strategies and a more complete understanding of the procedure for determining volume, as well as a facility at computation.

Problem 3: Shape A

Results

Percentage correct	Grade 4	Grade 8
Boys	24	70
Girls	25	57

This more abstract figure proved to be very difficult for most fourth graders. Miscounting the number of cubes accounted for the largest percentage of errors (24 percent of the students), closely followed by another 22 percent who confused area and volume (i.e., did not multiply all three dimensions). At the fourth grade level, there was no difference in the suc-



Shape A

cess rate of girls and boys. However, despite the similar success rate, there was a greater tendency for boys to treat the figure as a whole and multiply the dimensions, in contrast to girls who focused on individual missing cubes and counted (50 percent versus 39 percent, respectively).

The difference in strategy between boys and girls was magnified at grade eight. Most boys (61 percent) looked at the figure as a whole, computed the dimensions and then subtracted the existing cubes. This was the most successful strategy: 80 percent who followed it correctly calculated the number of cubes. However, it was used by only 42 percent of the girls. The others focused on the individual cubes needed to complete the figure and tried to count the missing cubes. As a result of this less efficient strategy, they were more susceptible to errors involved in miscounting or duplication.

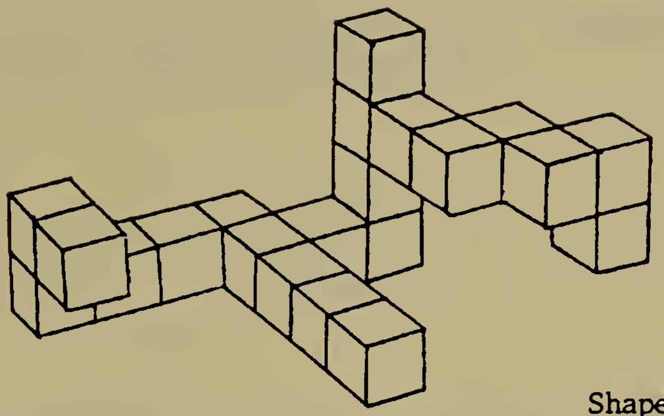
These results in regard to strategy, differing as they do in relation to age and gender, suggest that this problem may have called for abilities that differed from merely determining the volume of a box, as in the previous exercise. Whereas determining the capacity of the container could be regarded as a measurement exercise, the missing cubes in the figure may have called for visual imagery or experience in working with figures in three dimensions. Certainly, the tendency to view the figure globally as a rectangular solid helped in the solution of the problem.

Problem 4: Shape B

Results

Percentage correct of students completing problem	Grade 4 n=138	Grade 8 n=125
Boys	14	49
Girls	11	42

Approximately a third of the fourth grade students could not complete this problem and very few managed to calculate the correct number of missing cubes. Those that were successful adopted the strategy of calculating the total volume and subtracting the existing cubes to find those that were missing. However, the complexity of the figure made it difficult to determine just what those dimensions were.



Shape B

As was the case with the previous question, the lack of difference in success rate among boys and girls obscures the fact that there was a large difference in the kinds of errors that they committed. Exactly twice the percentage of boys than girls miscounted (56 percent versus 28 percent). Girls' errors tended to be more scattered, without a noticeable trend.

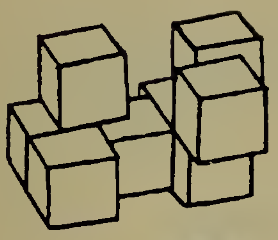
As was the case with the fourth graders, 97 percent of those eighth grade students who arrived at the correct solution did so by adopting the strategy of multiplying the dimensions and subtracting the missing cubes. The proportion of students whose errors could be attributed to miscounting was, at the eighth grade, higher among the girls (41 percent versus 29 percent of boys). This may have been due to the fact that 17 percent of the girls attempted to fill in layers and count the missing cubes, a strategy that was bound to lead to confusion.

Problem 5: Shape C

Results

Percentage correct among eighth grade students (n=99)	
Boys	28
Girls	29

Twenty eight percent of the eighth grade students either did not attempt this task or could not arrive at any solution. Generally, previously noted differences in strategy continued. Three-quarters of the boys measured the dimensions in contrast to half of the girls. More girls con-



Shape C

tinued to rely on their strategy of calculating the number of cubes in each layer and adding the missing cubes. Given the complexity of the figure, with its various missing half-cubes, miscounting caused most errors. Approximately 40 percent of the students made an error in counting, while 12 percent estimated incorrectly. Since boys preferred to multiply the dimensions, computation error was also a factor among boys.

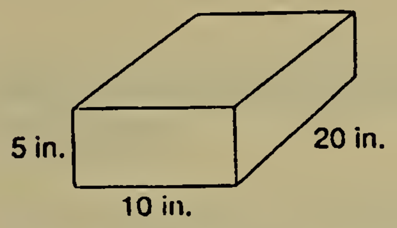
Summary: Overall Success

Percent of Students Giving Correct Solution to Problems

	Gr 4	Gr 8
none	45	17
Box only	29	21
Box and Shape A only	11	11
Box, Shapes A and B	3	13
Box, Shapes A, B, and C	na	11

For the vast majority of fourth graders, this set of problems was very challenging, despite the availability of concrete aids to help them calculate. Students who were able to multiply dimensions or count hidden cubes in a similar multiple-choice item succeeded in giving the correct answer. In addition, some girls profited from the use of the cubes by placing them in the bottom of the box and estimating how many layers of cubes it would take to fill it. However, the added burden of mentally constructing the necessary dimensions proved to be too difficult a task for most fourth graders and few completed Shapes A or B.

Three-quarter of eighth graders gave the correct answer when presented with the parallel multiple-choice question; half answered correctly to another question which gave the dimensions only (see below.) Why, then, were so few students able to calculate the missing cubes in this set of rectangular solids? We can only surmise that the requirement to work with actual figures rather than with numbers, to imagine how the solid would look when completed, and then to remember the requirements of the task (i.e., to give the *missing* number of cubes, not the total volume) was overwhelming to most students for whom such tasks were unfamiliar.



What is the volume of the rectangular solid above?

Percentages correct

	Grade 8
Boys	49
Girls	49

The Role of Strategy

Students' strategy differed according to their age, gender, and the demands of the problem itself. At the fourth grade level, there was little difference in approach between girls and boys. The majority looked at the shapes as a collection of cubes and tried to "fill in" those that were missing. When presented with the first block, Shape A, only 25 percent of the students adopted a strategy of multiplying the dimensions and subtracting the existing cubes. Perhaps because of their experience with the first shape, increasing confidence, or the demands presented by the shape itself, more looked at Shape B as a potentially completed solid and attempted a strategy of "multiplying the dimensions." Although few fourth graders were able to calculate the correct number for this shape, those who saw the figure in terms of its exterior dimensions were more successful. Their success may have been due to less opportunity for confusion in counting or estimating, or to those students' understanding of the concept of volume. Certainly, this was the more effective strategy.

Percentage of students adopting a "Multiply Dimensions" strategy and success at finding the solution.

	Grade 4		Grade 8	
	Boys	Girls	Boys	Girls
	"Multiplying Dimensions" Strategy(Success)		"Multiplying Dimensions" Strategy(Success)	
A	25(24)	25(25)	61(70)	42(57)
B	35(14)	33(11)	72(49)	65(42)
C	na	na	64(28)	38(29)

By eighth grade, differences in strategy began to emerge as more boys than girls treated the figure as a potentially completed solid, multiplying the dimensions and subtracting the existing cubes. Sixty-one percent of the boys, in contrast to forty-two percent of the girls, followed this strategy in calculating the cubes for Shape A. Clearly, it was more successful, as shown above. The analysis indicates that 80 percent of the students who followed this strategy were correct. When presented with Shape B, many girls changed strategies. They appeared either to learn with experience or to adapt their approach when faced with a shape that was unsuitable for filling in layers and counting (their preferred strategy for Shape A). Boys continued to use the "multiplying dimensions" strategy with the most difficult shape, Shape C, but girls did not. Only 38 percent of girls attempted to measure the exterior dimensions of Shape C. The majority reverted to their more familiar strategy of attempting to fill in the missing cubes and add. However, in the last two shapes, they were as successful as boys in the calculating the correct number of missing cubes.

It seems that, although boys and girls showed little difference in their success rate, older students did differ in how they went about the task. Furthermore, almost all girls who adopted the whole dimension strategy were successful (in contrast to many boys who were not despite their more efficient strategy). This suggests a distinction between the selection of a strategy and the execution. Many girls were effective, despite their less efficient strategy. This is an interesting result in view of the frequent contention that "girls are poor at spatial problems." In this case, when they had the opportunity to work with the cubes, their difficulty with spatial visualization (i.e., visualizing the incomplete solid as having a complete set of dimensions) did not bar them from calculating the correct answer. Experience, discussion about strategies, knowing how to look, and confidence may act as more important factors than some attribute associated with gender.

Attitudes

The problems were engrossing to students, particularly at the older level. Seventy-eight percent of fourth graders and 85 percent of eighth graders remained engaged throughout. Some attitudes showed greater variability, however, and were strongly related to how well students performed. For example, students who showed confidence and carefully checked their results were significantly more successful than others. Not only was confidence related to success, there was a significant difference in the confidence level of boys and girls. As shown in the table below, about 20 percent more boys than girls at each grade level were judged as confident by the administrators.

	Grade 4	Grade 8
Boys	57	66
Girls	36	42

Adaptation for Classroom Evaluation

The requirements for reliable results state-wide led to a very structured testing format. However, these tasks can be easily adapted to classroom testing. Using the students' charts of results, observation of their methods, and class or written discussion, information can be obtained on the level of students' understanding. Below is a list of behaviors that can be noted for each task, as well as supplementary questions that may be given orally or in written form.

Task 1 (Constructing Rectangular Solids):

From students' chart:

Number of solids recognized ___
Is recorded length always greater than width? Y/N

From observation:

Rejected 4 X 4 on grounds that it is a square Y/N
Attempted to use 3's or other non-factors of 16 Y/N
Recognize that solid is same despite orientation Y/N

Discussion questions (written or oral):

Do you see any kind of pattern on the sheet? Can you see how you could have come up with the shapes using only the chart? Why is this so?

What other sets of cubes could you use to make a set of rectangular solids? What sets of cubes *couldn't* you use? What is the difference between these sets?

Tasks 2 - 5: (Filling the Box and Completing the Solids.)

Observational Checklist:

Strategy:	multiplies dimensions	___
	adds/multiplies layers	___
	counts missing blocks	___
Result:	correct	___
	counting error	___
	estimation error	___
	confuses area/volume	___
	duplicates cubes	___
	does not subtract existing cubes from total	___
Attitude:	students work on tasks together	___
	one student dominates	___
	proceeds confidently toward solution	___
	remains engaged throughout	___
	careful with checking of results	___
	enthusiastic	___

Extension for Classroom Instruction

The tasks presented here can easily be used as a basis for extended investigations of factors, other number relationships, spatial visualization and the measurement of volume. They should be considered as opportunities for students to explore mathematics. The more connections that students can make between the visual and symbolic representation of mathematical concepts, the more confident they will feel in their understanding of number. Listed below are some related activities that can be used in the class. They are somewhat ordered by difficulty, but most can be adapted to any level.

Let students make their own boxes. They can use an existing rectangular solid to trace a network of faces, cut it out and tape it up. This can lead to a discussion of networks. Let them find for themselves which ones work and why.

If you use squared paper, students can explore the relationship between the surface area and the volume. Does the box that has the most squares on its faces hold the most cubes? Can you find two boxes that hold the same number of cubes but have different numbers of squares on their faces? Can you find two boxes that have the same number of squares on their faces but hold different numbers of cubes?

Using the 16 cubes, problems can be extended to look at other relationships. How many outside faces are there in each shape? How many interior faces? Is there a relationship between the two variables? How would you explain it? Does it work for solids of other volumes? Is there a way of predicting this for other rectangular solids?

Have the students construct different rectangular solids out of cubes, then paint them. How many cubes are painted on one face? On two faces? On three faces? Are not painted at all? What is the surface area of each solid? What is the volume? Now compare the results from different solids. Do any patterns form? Is it possible to make a prediction from the data?

designed by

Graphitopia

This pamphlet reports on the results of the 1989 testing carried out by the Massachusetts Educational Assessment Program. Other pamphlets in this series: "Animal/Leafkey," "Background Summary," "Circuits," "Insulation," "Math Town," "Playground," and "Popcorn Estimation."